

Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) EP 0 810 582 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
03.12.1997 Bulletin 1997/49

(51) Int Cl.⁶: G10L 5/02

(21) Application number: 97303442.4

(22) Date of filing: 20.05.1997

(84) Designated Contracting States:
DE FR GB

(72) Inventor: Miyashita, Shichiro
Yamato-shi, Kanagawa-ken (JP)

(30) Priority: 30.05.1996 JP 136327/96

(74) Representative: Zerbi, Guido Maria
Intellectual Property Department,
IBM United Kingdom Ltd.,
Hursley Park
Winchester, Hampshire SO21 2JN (GB)

(71) Applicant: International Business Machines
Corporation
Armonk, N.Y. 10504 (US)

(54) Voice synthesizing method, voice synthesizer and apparatus for and method of embodying a voice command into a sentence

(57) A method which is capable of performing an operation of creating a sentence and an operation of adjusting a voice attribute at the same time.

If a key for embedding an embedding command into an unsettled character string is pushed with the state where the unsettled character string has been displayed

after kana-kanji conversion, the voice attribute information held by a voice attribute information input section 115 is embedded in the form of an embedded command into the unsettled character string. Also, if a key for instructing voice synthesis is pushed with this state, voice synthesis is performed according to the embedded voice attribute information.

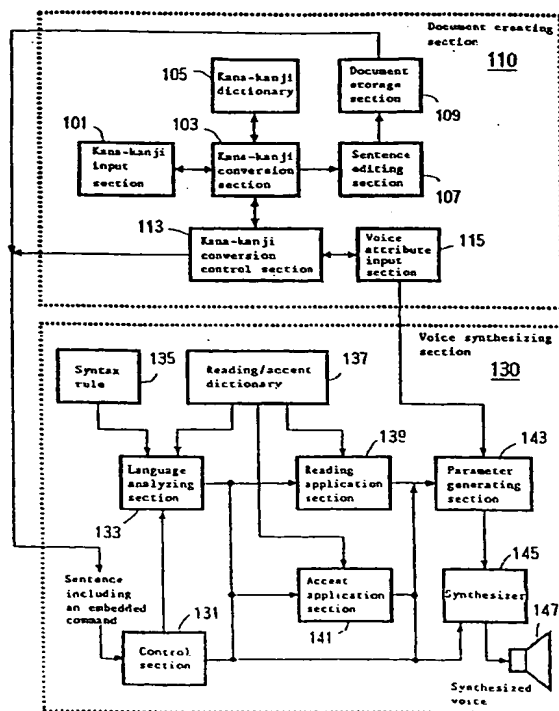


FIG. 2

Description

The present invention relates to voice synthesis, and more particularly to a method of creating a sentence embedded with a voice command which instructs a voice attribute for adjusting voice when voice synthesis is performed.

In many of conventional voice synthesizing programs, an operation of making a sentence for voice synthesis and an operation of adjusting voice synthesis are separately performed. For the operation of making a sentence for voice synthesis, first (1) a sentence is made by a kana-kanji conversion program, etc. Next, (2) the rough adjustment ("speed," "volume," etc.) of the entire system is performed. Finally, (3) words difficult to read are adjusted by using word registration, etc.

The voice synthesis is performed after the aforementioned operations (1) through (3) are all completed, and the voice synthesis cannot be performed while the voice adjustment is being performed. Also, it is general that the voice synthesis and the operations (2) and (3) are iterated by a cut-and-try method.

In the aforementioned voice synthesizing program, a sentence whose voice synthesis is desired is made once and the voice synthesis is performed. If the voice synthesis is unsatisfactory, an adjustment will be performed by the resetting of the volume of the entire system or the word registration, or by giving a reading attribute directly to the sentence. Thereafter, the voice synthesis is again performed and confirmed. However, these operations are respectively interrupted once and need to be iterated, so the operational efficiency is low.

In the case of ProTALKER/2 V1.0 which is one of the voice synthesizing programs, in addition to the functions that the aforementioned general voice synthesizing programs have, there are the following features (a) and (b):

(a) A command, which changes an attribute that only a program can interpret, can be embedded as an embedded command into a sentence whose voice synthesis is performed. After this command, the voice synthesis of the sentence is performed by a specified attribute until the next command appears. The embedded command can set "distinction of sex," "speed," "volume," "pitch," "intonation," and so on. Since "reading," "accent," etc., are not supported in units of a word by the embedded command, they can not be registered temporarily as word registration.

(b) The embedded command is assumed to be input with keys by users.

In the case of "ELOQUENT SPEAKER" which is one of the voice synthesizing programs, in addition to the functions that the aforementioned general voice synthesizing programs have, there are the following fea-

tures (a), (b), and (c):

(a) On a special editing window which is opened while a sentence for voice synthesis is being made, not only "reading" and "accent" but also "accent strength," "breathing-pause length" at a place of breathing-pause, "volume," and "speed" can be adjusted in units of articulation.

(b) "Reading" of each articulation at fine setting can be selected from an all-candidate panel, and users do need to input it. However, for other attributes ("accent strength," "breathing-pause length," etc.), users need to directly input them as in the case of ProTALKER/2.

(c) The fine attributes with respect to the sentence set at (a) and (b) are stored as an attribute file. When the voice synthesis of the sentence is performed, the attribute file, together with the sentence file, is read in and utilized.

Even in the aforementioned voice synthesizing programs "ProTALKER/2" and "ELOQUENT SPEAKER," an operation of creating a sentence and an operation of adjusting a voice attribute cannot be performed at the same time. Therefore, after a whole sentence is created, the entire sentence or character string specified in the document need to be input to perform voice synthesis. As compared with the case where a sentence is created while voice is being confirmed, the operational efficiency is low and consequently, these synthesizing programs are unsuitable to make a voice-command embedded sentence in a short time. In addition, in these methods, attribute commands need to be input directly with keys by users, and consequently, memorizing or looking over attribute commands of various kinds becomes troublesome as it is complicated. Furthermore, there is a possibility of a mistaken input, because a key input must be performed directly.

On the other hand, in Published Unexamined Patent Application No. 5-143278 there is disclosed a method which performs voice synthesis in correspondence with the style of type (Ming type, Gothic type, etc.), emphasis (full angle, half angle, etc.), and decoration (underline, netting, etc.) of a character string existing in a document. In such a method, it is unclear how a character string where the style of type, the emphasis or the decoration was changed is synthesized to voice which has what kind of attribute, and a great deal of skill is required. In addition, this method does not give suggestions as to how the voice synthesis of only a character string where the style of type was changed is performed, and consequently, the entire document needs to be input to perform voice synthesis.

Also, in Published Unexamined Patent Application No. 6-176023 there is disclosed a method where the voice synthesis of a character string existing in a docu-

ment is performed with priority given to the reading of a kana (Japanese character) which is input at the time of kana-kanji conversion. For example, when a character string "market (Japanese kanji for market has two readings: "ichiba" and "shijo")" is obtained by inputting "ichiba (kana)" rather than "shijo (kana)" and converting the kana to the kanji ("market" in this case), the voice synthesis of the "market" is performed as "ichiba." This method can change the reading of a kanji only when it has two or more readings, however, it is impossible to change the voice attribute of a character string in a manner desired by a user. Also, this method changes the priority of a reading-accent dictionary which is used when performing the voice synthesis. Therefore, once word registration is performed so that the "market" in a certain sentence is pronounced as "ichiba," the market will be pronounced as "ichiba" even in other sentences where it is desired that the "market" is pronounced as "shijo."

It is an object of the present invention to provide a technique which alleviates the above drawbacks.

According to the present invention we provide a method of creating a sentence embedded with a voice command which includes voice attribute information and which is referred to when voice synthesis is performed, the method comprising the steps of: specifying a character string into which said voice command is embedded; detecting a user's input which instructs embedding of said voice command into said specified character string; displaying entries for the user to input voice attribute information of said specified character string; and embedding a voice command, which includes voice attribute information corresponding to the user's input to said entries, into said specified character string.

Further according to the present invention we provide an apparatus for creating a sentence embedded with a voice command which includes voice attribute information and which is referred to when voice synthesis is performed, the apparatus comprising: an unconverted character string input section for holding a character string input by a user; a character conversion dictionary for managing a converted character string which corresponds to an unconverted character string; a character conversion section for retrieving a candidate for a converted character string which corresponds to said character string held by said unconverted character string input section; a voice attribute input section for holding a voice attribute value adjusted by a user's input; and a character conversion section for instructing said character conversion section to select a converted character string corresponding to the character string held by said unconverted character string input section from said converted character string candidate in response to a user's input and also for embedding said voice attribute value held by said voice attribute input section into the converted character string selected in the form of a voice command.

According to a preferred embodiment of the present

invention, a function of embedding an embedded command into an unsettled character string is allocated to a certain key, and if the key is pushed, the unsettled character string will be converted to an unsettled character string embedded with the command. Also, if a key instructing voice synthesis is pushed with the state where the unsettled character string has been displayed after kana-kanji conversion, voice synthesis will be performed according to the reading attribute valid at that time, and at the same time, the unsettled character string will be converted to the format where embedded commands representative of attributes have been added. Then, for example, by changing the attributes by using a control panel, voice synthesis can be performed many times at that place. Also, the unsettle character string is suitably changed according to the attribute at that time. Furthermore, in the case where a plurality of articulations (conversion object character string) exist in a single unsettled character string and where it is desired that a certain articulation and the articulations thereafter are read at a different attribute, a cursor is moved to that articulation and after the attribute of the articulation is again adjusted, an embedded command can be embedded before the articulation by pushing a key for this voice synthesis. In this way, the certain articulation and the articulations thereafter are read at the adjusted attribute.

A function of starting word registration valid only temporarily is allocated to a certain key, and a word for which word registration is desired is segmented in units of articulation. If the key is pushed with the state where the word can be converted, the function of the word registration which is valid only temporarily will be called out with the word as a word to be registered. It is preferable that a user interface be nearly identical with ordinary word registration, and registered information is not registered in a user dictionary but is embedded into an unsettled character string as an embedded command. A quantity of information to be embedded is matched with that of information which is registered in ordinary word registration. Then, if a settling key is pushed by a user, a character string into which the embedded command was inserted will be sent to an editing application. At this point, voice synthesis can also be performed again.

In a preferred embodiment of the present invention, there is provided a method of creating a sentence embedded with a voice command which is referred to when voice synthesis is performed, comprising the steps of: holding a kana character string input from the input unit in the character string input section as an unsettled character string; detecting a user's input, which instructs conversion to a kanji-kana mixed character string with respect to the unsettled character string input, from the input unit; specifying a candidate character string, which is a candidate for a kanji-kana mixed character string corresponding to a conversion object character string forming part of the unsettled character string, from the kana-kanji dictionary in response to the detection of the

input which instructs conversion to a kanji-kana mixed character string; displaying the candidate character string on the display; detecting a user's input, which selects a selected character string which is one of the candidate character strings, from the input unit; replacing the conversion object character string with the selected character string and taking the selected character string to be a new unsettled character string; detecting a user's input which instructs embedding of the voice command into the conversion object character string; displaying entries for the user to input voice attribute information which includes reading and accent of the conversion object character string which are embedded into the conversion object character string; embedding a voice command, which includes voice attribute information corresponding to the user's input to the entries, into the conversion object character string; detecting a user's input which instructs voice synthesis of the conversion object character string; and performing voice synthesis in accordance with a voice attribute of the voice command.

In another preferred embodiment of the present invention, there is provided a method of creating a sentence embedded with a voice command which is referred to when voice synthesis is performed, comprising the steps of: holding a kana character string input from the input unit in the character string input section as an unsettled character string; detecting a user's input, which instructs conversion to a kanji-kana mixed character string with respect to the unsettled character string input, from the input unit; specifying a candidate character string, which is a candidate for a kanji-kana mixed character string corresponding to a conversion object character string forming part of the unsettled character string, from the kana-kanji dictionary in response to the detection of the input which instructs conversion to a kanji-kana mixed character string; displaying the candidate character string on the display; detecting a user's input, which selects a selected character string which is one of the candidate character strings; from the input unit; replacing the conversion object character string with the selected character string and taking the selected character string to be a new unsettled character string; detecting a user's input which instructs embedding of the voice command into the conversion object character string; displaying entries for the user to input voice attribute information which includes reading and accent of the conversion object character string which are embedded into the conversion object character string; and embedding a voice command, which includes voice attribute information corresponding to the user's input to the entries, into the conversion object character string.

In another preferred embodiment of the present invention, there is provided an apparatus for creating a sentence embedded with a voice command which is referred to when voice synthesis is performed, comprising: a kana character string input section for holding a character string input by a user; a kana-kanji dictionary

for managing a kanji-kana mixed character string which corresponds to a kana character string; a kana-kanji conversion section for retrieving a candidate for a kanji-kana mixed character string which corresponds to the character string held by the kana character string input section; a voice attribute input section for holding a voice attribute value adjusted by a user's input; and a kana-kanji conversion section for instructing the kana-kanji conversion section to select a kanji-kana mixed character string corresponding to the character string held by the kana character string input section from the kanji-kana mixed character string candidate in response to a user's input and also for embedding the voice attribute value held by the voice attribute input section into the kanji-kana mixed character string selected in the form of a voice command.

In another preferred embodiment of the present invention, there is provided an apparatus including a document creating section for creating a sentence embedded with a voice command which includes voice attribute information and which is referred to when voice synthesis is performed, also including a parameter generating section for generating parameters which are used for voice synthesis, and further including a voice synthesizing section for performing voice synthesis from an input sentence, the apparatus comprising: a character string input section for holding a character string input by a user; a voice attribute input section for holding a character string voice attribute value which instructs reading of the character string adjusted by a user's input; a conversion control section for embedding the character string voice attribute value held by the voice attribute input section into the character string input in the form of a character string voice command in response to a user's input; and a voice synthesis control section for instructing the parameter generating section to perform voice synthesis in accordance with character string voice attribute information embedded in the character string embedded with the character string voice command.

In another preferred embodiment of the present invention, there is provided an apparatus for performing voice synthesis of a sentence which includes voice attribute information, comprising: a kana character string input section for holding a character string into which a voice command is embedded; a kana-kanji dictionary for managing a kanji-kana mixed character string which corresponds to a kana character string; a kana-kanji conversion section for retrieving a candidate for a kanji-kana mixed character string which corresponds to the character string held by the kana character string input section; a voice attribute input section for holding a voice attribute value adjusted by a user's input; a kana-kanji conversion section for instructing the kana-kanji conversion section to select a kanji-kana mixed character string corresponding to the character string held by the kana character string input section from the kanji-kana mixed character string candidate in response to a user's

input and also for embedding the voice attribute value held by the voice attribute input section into the kanji-kana mixed character string selected in the form of a voice command; and a voice synthesizing section for performing voice synthesis in accordance with voice attribute information embedded in the kanji-kana mixed character string embedded with the voice command.

In another preferred embodiment of the present invention, there is provided an apparatus for performing voice synthesis of an input sentence, comprising: a language analyzing section for determining reading and accent of a character string which is included in the input sentence, based on syntax rule information and a reading/accent dictionary; a voice synthesizing unit for performing voice synthesis in accordance with the reading and accent of the character string which is included in the input sentence, determined by the language analyzing section; and a voice synthesis control section which, when there is embedded a voice command which corresponds to the input character string and also instructs a voice attribute value of a vice attribute including reading and accent of the input character string when voice synthesis is performed, performs voice synthesis of the character string in accordance with the voice attribute value instructed by the voice command.

A preferred embodiment of the present invention will hereinafter be described in reference to the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram showing hardware constitution;

Figure 2 is a block diagram of processing elements;

Figure 3 is a diagram showing a user interface of the present invention;

Figure 4 is a diagram showing an embedded sentence command of the present invention;

Figure 5 is a diagram showing a user interface of the present invention;

Figure 6 is a diagram showing an embedded character string command of the present invention;

Figure 7 is a flowchart showing a procedure of creating a sentence which includes an embedded command of the present invention;

Figure 8 is a flowchart showing a procedure of creating a sentence which includes an embedded command of the present invention;

Figure 9 is a flowchart showing the control procedure that is performed by a voice synthesis control

section which received a sentence including an embedded command of the present invention; and

Figure 10 is a diagram showing a user interface of the present invention.

Referring to Figure 1, there is shown a block diagram of hardware constitution for carrying out a voice synthesizing system of the present invention. The voice synthesizing system 100 includes a central processing unit (CPU) 1 and a memory 4. The CPU 1 and the memory 4 are connected to a hard-disk drive 13 serving as a secondary storage through a bus 2. A floppy-disk drive (or a disk drive for an magneto-optical (MO) memory or a compact disk read-only memory (CD-ROM)) 20 is connected to the bus 2 through a floppy-disk controller 19.

Inserted into the floppy-disk drive (or a disk drive for an MO memory or a CD-ROM) 20 is a floppy disk (or a recording medium such as an MO memory or a CD-ROM). The floppy disk, the hard-disk drive 13, and ROM 14 can give an instruction to the CPU in cooperation with an operating system and record the codes of a computer program for implementing the present invention. The codes can be executed by loading them into the memory 4. The codes of this computer program can be compressed, or they can be segmented into a plurality of parts and recorded on a plurality of recording media.

The voice synthesizing system 100 can be further made a system equipped with user interface hardware.

The user interface hardware includes, for example, a pointing device (such as a mouse and a joy stick) 7 or keyboard 6 for inputting data and a display 12 for presenting visual data to users. It is also possible to connect a printer through a parallel port 16 or to connect a modem through a serial port 15. Furthermore, it is possible for the voice synthesizing system 100 to communicate with another computer through the serial port 15 and the modem, or through a communication adapter 18. A speaker 23 receives a voice signal supplied from an audio controller through an amplifier 22 and outputs the signal as voice. Thus, it easily follows that the present invention is executable by general personal computers (PCs) or work stations (WSs). Note that the aforementioned constituents are examples and that all of the constituents do not always become the requisite elements of the present invention.

It is desirable that the operating system of the present inventions be an operating system, such as Windows (Microsoft trademark), an OS/2 (IBM trademark), and an X-WINDOW system on AIX (IBM trademark), which supports a GUI multi-window environment at standard. The present invention, however, is executable even under a character-based environment such as PC-DOS (IBM trademark) and MS-DOS (Microsoft trademark) and is not limited to a specific operating system's environment. Although Figure 1 shows a stand-alone system, the present invention may be realized as a client/server system. A client machine may be con-

nected to a server machine through an internet or through a local area network (LAN) by a token ring. On the side of the client machine, only a kana character string input section forming part of a document generating section to be described later, a synthesizer for receiving voice data from the sever machine side and re-constituting it, and a speaker may be disposed, while the other functions may be disposed on the sever machine side. Thus, it is a freely changeable design matter what functions are disposed on the server machine side and the client machine side, and various modifications, such as what functions are disposed and executed to a combination of machines, are concepts which are included within the ideas of the present invention.

B. SYSTEM CONFIGURATION

The system constitution of the present invention will next be described in reference to a block diagram of Figure 2. A preferred embodiment of the present invention is roughly constituted by a document creating section 110 and a voice synthesizing section 120. The document creating section 110 and the voice synthesizing section 120 can be separately realized by the hardware constitution shown in Figure 1 or they can be realized by shared hardware.

The document creating section 110, as is shown in the figure, is constituted by a kana character string input section 101, a kana-kanji conversion section 103, a kana-kanji dictionary 105, a document editing section 107, a document storage section 109, a kana-kanji conversion control section 113, and a voice attribute input section 115.

The document creating section 110 creates and stores a sentence embedded with an embedded command which becomes an input for voice synthesis. The kana character string input section 101 holds an input signal, input from the keyboard 6, as an unsettled character string. In a preferred embodiment of the present invention, a buffer which manages kana-kanji conversion software corresponds to this kana character string input section. In a preferred embodiment of the present invention, while the present invention has been carried out by improving kana-kanji conversion software, the ideas of the present invention are not limited to this. For example, for the character string of a sentence which has already been settled, a range can be specified to specify a character string by using the pointer of the mouse 7, or the like, and the specified character string can be copied to a buffer which is managed by the kana character input section 101. In such a case, after the conversion of the present invention to be described later is performed, the specified character string in the settled document is deleted, or immediately before the character string, the converted character string is put into.

The kana character string conversion section 103 retrieves the kana-kanji dictionary 105 to convert the unsettled character string to a kanji-kana mixed character

string which corresponds to the character string held by the kana character string input section 101. The kana-kanji dictionary 105 stores a kanji-kana mixed character string corresponding to a kana character string, and the kana character string conversion section 103 retrieves a kanji-kana mixed character string corresponding to an unsettled character string. At this time, there are cases where an unsettled character string is longer than a character string with a length corresponding to the character string held by the kana-kanji dictionary. In such a case, preferably a morphological analysis is performed and the unsettled character string is divided so as to correspond to the length of the character string held by the kana-kanji dictionary. The character string, where the division is performed and which becomes an object of conversion by pressing a present conversion key, is called the conversion object character string. In the case where a kana character string is converted to a kanji-kana mixed character string, the conversion is processed in units of the conversion object character string. Preferably, this conversion is displayed in the display screen in the format which can be distinguished from an unsettled character string (for example, in an unsettled character string, the part of the conversion object character string is displayed in a reversed manner and the remaining parts of the unsettled character string are displayed with underlines).

There are also cases where a plurality of kanji-kana character strings corresponding to a kana character string exist. In a preferred embodiment of the present invention, when a plurality of kanji-kana character strings exist like this, each character string (candidate character string) is given the priority order and displayed in a display unit in accordance with the priority order. Users can select a desired kanji-kana mixed character string from the kanji-kana mixed character strings which become candidates for the aforementioned conversion. By this user's selection, the unsettled character string held by the kana character input section 101 is replaced with the kanji-kana mixed character string selected by the user.

The sentence editing section 107 receives a kanji-kana mixed character string from the kana-kanji conversion section 103 and edits the character string. In a preferred embodiment of the present invention, the sentence editing section 107 corresponds to word processing software. The document storage section 109 stores the edited result of the sentence editing section in a recording medium.

The kana-kanji conversion control section 113 determines by the input instructed by a user (for example, input of a "conversion key" or a numeral value) which kanji-kana mixed character string is adopted among the kanji-kana mixed character candidates corresponding to the character string held by the kana character string input section, and instructs the kana-kanji conversion section to perform conversion. In the present invention, the kana-kanji conversion control section 113 also has

a function of embedding a voice attribute embedding command which instructs voice attribute change, based on the contents of the voice attribute adjustment entries adjusted by a user, when voice synthesis is performed.

The voice attribute input section 115 holds a user's input which instructs voice attribute change. The voice attribute input section will be described in detail later. The data held by the voice attribute input section is put into an unsettled character string or a conversion object character string, but, preferably it is possible, for example, to instruct the voice synthesizing section 130 to change the voice attribute of the default in voice synthesis by using the voice attribute input section 115. In such a case, the parameter information, managed by a parameter generating section 143 and a synthesizer 145 which are described later, is updated (for example, in the case of a voice attribute "volume," the synthesizer 145 can be instructed to raise the volume of a synthesized voice, and in the case of a voice attribute "intonation," the parameter generating section 143 can be instructed to change parameters). The voice attribute input section 115 is disposed in the document creating section 110, but it can also be included in the voice synthesizing section 130. The voice attribute input section 115 may be disposed in both the document creating section 110 and the voice synthesizing section 130 so that updated voice attribute data can be transmitted therebetween.

On the other hand, the voice synthesizing section 130 is constituted by a voice synthesis control section 131, a language analyzing section 133, a syntax rule holding section 135, a reading-accent dictionary 137, a reading application section 139, an accent application section 141, a parameter generating section 143, a voice synthesizing section 145, and a voice generating section 147.

The voice synthesis control section 131 receives the command embedded sentence stored in the document storage section 109 of the document creating section 110 or the command embedded character string transmitted from the kana-kanji conversion control section 113 of the document creating section 110. Based on the embedded command, the voice synthesis control section 131 discriminates a character string where reading and accent have been instructed and a character string where reading and accent have not been instructed from each other. The voice synthesis control section 131 sends the instructed character string to the language analyzing section 133 and the uninstructed character string directly to the parameter generating section 143. When an embedded command instructing parameter change is detected, the parameter change is instructed to the parameter generating section 143.

Note that it is also possible that the voice synthesis control section 131 sends not only the instructed character string but also the uninstructed character string to the language analyzing section 133. In such a case, the reading and accent determined by the language analyzing

section 133 is ignored and the reading and accent instructed by the embedded command are prior. In this method, in order to match the character string segmentation instructed by an embedded command with the character string segmentation performed by the language analyzing section 133, it is desirable that a delimiter or command instructing the segmentation instructed by an embedded command be sent to the language analyzing section 133.

The language analyzing section 133 performs the morphological analysis of the character string transmitted from the voice synthesis control section 131 by referring to both the reading/accent dictionary 137 and the syntax rule stored in the syntax rule holding section 135, and the language analyzing section 133 segments an input sentence into appropriate morphological units.

The syntax rule storage section 135 stores syntax rules which are referred to in the morphological analysis in the language analyzing section 133. The reading-accent dictionary 137 stores "a part of speech," "reading," and "accent" which correspond to a kanji-kana mixed character string.

The reading application section 139 determines the readings of the individual morphemes segmented by the language analyzing section 133 from the reading information stored in the reading-accent dictionary 137.

The accent application section 141 determines the accents of the individual morphemes segmented by the language analyzing section 133 from the accent information stored in the reading-accent dictionary 137.

The parameter generating section 143 generates voice parameters for performing voice synthesis with currently specified parameters, such as speed, pitch, volume, intonation, and distinction of sex, in accordance with the reading determined by the reading application section 139 and the accent determined by the accent application section 141. What is meant by the "currently specified parameters" is that when a voice command representative of a voice attribute is embedded before the character string where the voice synthesis is presently being performed, the voice attribute is adopted and that when there is no such a command, the voice attribute value of the default previously set in the system is adopted.

The voice synthesizer 145 generates a voice signal in accordance with the voice parameters generated by the parameter generating section 143. In a preferred embodiment of the present invention, the generation of the voice signal is performed by performing digital/analog (D/A) conversion by means of the audio controller of Figure 1. In accordance with the voice signal generated by the voice synthesizer 145, the voice generating section 147 generates voice. In a preferred embodiment of the present invention, the generation of the voice is performed by the amplifier 22 and speaker 23 of Figure 1.

While the functional blocks shown in Figure 2 have been described, these functional blocks are logic functional blocks and it is meant that the functional blocks

are realized not by respective hardware and software but by composite or shared hardware and software.

Figures 7 and 8 are flowcharts showing a preferred embodiment of the present invention. First, the kana-kanji conversion control section 113 of the document creating section 110 of the present invention judges whether there is an unsettled character string or not (step 404). In a preferred embodiment of the present invention, the judgment of whether there is an unsettled character string or not is performed based on whether data exists in the buffer managed by the kana-kanji conversion control section 113 or not. By inputting characters through the keyboard 6 during operation of kana-kanji conversion software, data is accumulated in the buffer managed by the kana-kanji conversion control section. When an unsettled character string does not exist, the kana-kanji conversion control section 113 waits for an unsettled character string until it is input. When an unsettled character string exists, the unsettled character string is displayed (step 405). In a preferred embodiment of the present invention, an unsettled character string is settled, and in order to distinguish a settled character string and an unsettled character string sent to the editing section 107, an unsettled character string is emphatically displayed with underlines or inverted display.

With the state where the unsettled character string exists, the kana-kanji conversion control section 113 waits until any key is pushed (step 407). When the input key is a kana-kanji conversion key (step 409), the kana-kanji conversion control section 103 selects a kanji-kana mixed character string having the highest priority order or a kanji-kana mixed character string selected by a user from the kana-kanji dictionary 105, and the selected character string is taken to be a new unsettled character string (step 411). That is, the content of the buffer managed by the kana-kanji conversion control section 113 is replaced with this character string.

Next, when the input key is a voice synthesis key (step 413), the voice attribute information at that time is acquired (step 415). In a preferred embodiment of the present invention, a specific PF key is allocated as the voice synthesis key, and the kana-kanji conversion control section 113 will judge that the voice synthesis key has been pushed, if the PF key is input. However, the voice synthesis key is not limited to the PF key, but may be a specific key or a combination of keys of the keyboard 6, or may be a button icon which instructs the embedding of a voice synthesis command specified by the mouse 7. What is meant by the "voice attribute information at that time" is that in a preferred embodiment of the present invention the attribute information of the default exists and also in the case where any voice attribute information about the sentence is not defined, voice synthesis is performed according to the attribute information of the default. In a preferred embodiment of the present invention, a panel 303 is provided for changing voice attribute information and voice attributes can be

defined by entries 311 through 329 for changing each voice attribute information on the panel 303.

As shown in Figure 3, the panel 303 includes entries 311 and 313 for changing "speed" which is one of the voice attributes, entries 315 and 317 for changing "pitch," entries 319 and 321 for changing "volume," entries 323 and 325 for changing "speed," and entries 327 and 329 for changing "distinction of sex." In a preferred embodiment of the present invention, the default values of the voice attributes have previously been set in the system, and when a user does not change a voice attribute value, the voice attribute is displayed at the default value. When a user changes a voice attribute value, the voice attribute is displayed at the last voice attribute value changed.

A user can perform, for example, the adjustment of the speed which performs voice synthesis by dragging a slider 311 with the pointer of the mouse or the like. The speed can also be adjusted by inputting a numerical value directly to an attribute input portion 313. In a preferred embodiment of the present invention, as sliders 311, 315, 319, and 323 are changed, the numerical values of attribute value input portions 313, 317, 321, and 325 are also changed and displayed. Conversely, as the numerical values of the attribute value input portions 313, 317, 321, and 325 are changed, the sliders 311, 315, 319, and 323 are also changed and displayed. Also, the voice attribute, distinction of sex, can be specified by clicking on the entries 327 and 329 for changing distinction of sex.

In a preferred embodiment of the present invention, the present invention has been realized by the operating system which supports a GUI multi-window environment at standard, however, the present invention is executable under a character based environment which does not support the GUI multi-window environment. In such a case, entries for inputting voice attribute values as numerical values or characters are provided to users. The entries for adjusting voice attributes, shown in Figure 3, are examples, and having all of the voice attributes shown here as voice attributes is not the requirement of the present invention. In addition, other attributes, such as breathing-pause length, may be included. Furthermore, the entries for adjusting voice attributes are matters which are changeable at a stage of design, and all of such various changes are concepts included within the ideas of the present invention.

Next, if a button icon 331 for "O.K." shown in Figure 3 is pushed by a user after the adjustment of the voice attribute (step 417), the adjusted voice attribute value will be embedded in the form of an embedded command into the unsettled character string (step 419). In a preferred embodiment of the present invention, an embedded sentence command which is embedded into a sentence has been embedded in the format shown in Figure 4. In the figure, the embedded command starts with "[*" and ends with "]*". Also, "ASU ha HARE de sho (It will be fine tomorrow)" indicates an unsettled character

string. The "ASU" used herein is intended to mean a Japanese kanji which corresponds to "tomorrow." The voice synthesizing section 130 can identify a symbol representative of the start of this embedded command and a symbol representative of the end of the embedded command and thereby can discriminate the embedded command from an ordinary character string. Explaining the contents of this embedded command, the "M" of "[*MS9P81G8Y3]" indicates that the voice attribute of distinction of sex is a male. In the case "F", it indicates a female. "S9" indicates that "speed" is 9. "P81" indicates that "pitch" is 81. "G8" indicates that "volume" is 8. Finally, "Y3" indicates that "intonation" is 3.

The aforementioned method of embedding a symbol indicating the kind of a voice attribute and a value of the voice attribute as a set into a voice command is merely an example. The voice command may be embedded in a method where the voice synthesis control section 131 of the voice synthesizing section 130 can judge the voice command, the kind of the voice attribute embedded in the voice command, the value of the voice attribute, and the position in a sentence where voice attribute change is performed. For example, the voice attributes may be fixedly set so that the first byte of the voice command is "distinction of sex," the second byte is "speed," and so on, and the voice synthesis control section 131 may judge the kind of the voice attribute in accordance with the position in the voice command. Also, it is preferable that an embedded command be embedded in the head of a character string which validates the voice attribute included in the command. However, if the position in a sentence of a character string which validates the voice attribute is known, the command does not need to be embedded in the head of the character string. In this case, the position in a sentence of a character string, which validates a voice attribute embedded in a voice command, can be embedded at the voice command, and when voice synthesis is performed, the voice synthesis control section 131 can validate the voice attribute of the voice command when located at the position in the sentence of the character string which validates the voice attribute embedded in the voice command.

Next, the unsettled character string embedded with the aforementioned command is held as a new unsettled character string by the kana character string input section 101. However, the embedding of the embedded command may be performed not by pushing the O.K. button but by pushing a confirmation button to be described later. When an embedded command is embedded by the confirmation button, the voice attribute of the voice attribute entry in the final state changed by a user is embedded as a voice command. Note that, in response to this confirmation button being pushed, the present unsettled character string with the embedded command can also be sent to the voice synthesizing section 130 (Figure 2) to perform voice synthesis.

When a button icon 333 for "deletion" of Figure 3 is

selected, the embedded command of the present unsettled character string is deleted. Therefore, if a settling key to be described later is pushed in that state, the voice synthesis of the character string will be performed according to the attribute information at that time.

In the case where a button icon 335 for "voice synthesis" is pushed, when the voice attribute information of the unsettled character string has been changed, the unsettled character string is sent to the voice synthesizing section 130 with the state where the embedded command has been embedded in the unsettled character string, and the voice synthesis is performed. On the other hand, when the voice attribute information of the unsettled character string has not been changed, the voice attribute information at that time is embedded in the form of an embedded command and sent to the voice synthesizing section 130, in which the voice synthesis is performed. In a preferred embodiment of the present invention, the "voice attribute information at that time" has been temporarily stored, and an embedded command is created from the temporarily stored information. However, in the case of the default state, the embedding of an embedded command is not performed and an unsettled character string with no embedded command is sent to the voice synthesizing section 130. The parameter generating section 143 generates a voice parameter which has the previously set default value.

Next, when an input key is a temporary word registration key (step 427), a temporary word registration panel 305 shown in Figure 5 is opened (step 429). In this example, in the unsettled character string "ASU ha HARE de sho (It will be fine tomorrow)," the conversion object character string ASU (a kanji corresponding to "tomorrow"), which is the conversion unit of kana-kanji conversion, has been specified as a conversion object. With this state, the temporary word registration key is pushed and entries are displayed on the temporary word registration panel 305 for adjusting the voice attribute information of the character string "ASU." The temporary word registration panel 305 is provided with entries 343 and 347 for adjusting "accent," an entry 345 for adjusting "reading," and an entry 349 for adjusting "a part of speech." Users can apply a desired accent or reading to the "ASU." For example, the "ASU (a kanji corresponding to "tomorrow")" can be pronounced not as "asu (a kana corresponding to the kanji ASU)" but as "myonichi (a kana corresponding to the kanji ASU)," or accent different from ordinary accent can be specified.

Now, in the case where the button icon 355 for voice output is pushed (step 431), when temporarily registered information, such as "reading," "accent," and "a part of speech," exists, the character string voice attribute information is embedded in the form of an embedded command into a conversion object character string. The conversion object character string with the embedded command is sent to the voice synthesizing section 130, and the voice synthesis is performed (step 433). On the other hand, when temporarily registered

information, such as "reading," "accent," and "a part of speech," does not exist, a conversion object character string, as it is, is sent to the voice synthesizing section 130, and the voice synthesis is performed. In this case, the conversion object character string having no voice attribute information is given "reading" and "accent" by using the voice synthesizing section 130, the syntax rule 135, and the reading/accent dictionary 137.

When the "O.K." button icon 351 is pushed (step 435), character string voice attribute information, such as temporarily registered "reading," "accent," and "a part of speech," is embedded in the form of an embedded command, and the command embedded character string is taken to be a new unsettled character string (step 437). A preferred example of the character string embedded with this character string voice attribute information is shown in Figure 6.

Explaining the contents of the aforementioned embedded command, the "[*T" of "[*T asu ASU 0 000020 OB 1800]" is a symbol indicating the start of an embedded command of temporary word registration (the start of a character string voice command). As previously described, the "asu" is a kana corresponding to "tomorrow," and the "ASU" is a kanji corresponding to "tomorrow." The voice control section 131 of the voice synthesizing section 130 can judge the character string voice attribute embedded in the character string voice command by detecting the symbol, [*T.

The "asu" of the aforementioned character string voice command "[*T asu ASU 0 000020 OB 1800]" indicates the reading of the conversion object character string which validates the voice attribute information included in the character string voice command. The "ASU" specifies the conversion object character string included in the character string voice command. The voice synthesis control section 131 of the voice synthesizing section 130 stops sending the character string specified by the character string voice command to the language analyzing section 133 and directly instructs the parameter generating section 143 to generate voice synthesis parameters and the synthesizer 145 to perform voice synthesis. In a preferred embodiment of the present invention, the voice synthesis control section 133 judges the contents of the voice command and directly instructs the parameter generating section 143 and the synthesizer 145 to generate voice synthesis parameters and perform voice synthesis. However, it is also possible to perform a desired voice synthesis by giving information to the reading application section 139 and the accent application section 141.

The "0" of the embedded command "[*T asu ASU 0 000020 OB 1800]" is a voice attribute value indicating the position of accent, and the "000020" is information about a part of speech and is voice attribute information indicating information such as a proper noun and a gerund. The "OB" is a type and is voice attribute information indicating information such as a suffix, a prefix, and a general word. The "1800" is additional information and

is, for example, voice attribute information indicating additional information such as whether there is the nature attached to the prefix. Finally, the "]" is a symbol indicating the end of the voice command.

In a preferred embodiment of the present invention, the conversion object character string "ASU" is converted to a character string where a character string voice command is embedded before the conversion object character string, as in "[*T asu ASU 0 000020 OB 1800] ASU". However, for example, the conversion object character string may be converted to a string where a character string voice command and a symbol indicating the end of a command are embedded before and after the conversion object character string, as in "@asu@ 0 000020 OB 1800 ASU". Such a matter can be changed in various ways at the stage of design.

In a preferred embodiment of the present invention, the order of the voice attributes included in the character string voice command has been determined. By partitioning off the voice attribute by a delimiter (a space character), the voice synthesis control section 131 can judge the voice attribute included in the character string voice command. However, even in this character string voice command, as with the sentence voice command, the form of the voice attribute command shown here is merely an example, and consequently various changes are possible.

Referring again to Figure 8, in the case where the button icon 353 for deletion is pushed (step 439), a conversion object character string including an embedded command is replaced with a conversion object character string including no embedded command.

Next, when an input key is a settling key (step 451), an unsettled character string is sent to the sentence editing section 107 as a settled character string (step 455). Therefore, a character string having an embedded command with sentence voice attribute information or character string voice attribute information is sent to the sentence editing section 107 as a settled character string. Therefore, in the example of Figures 4 and 6, a settled character string such as "[*MS9P81G8Y3] [*T asu ASU 0 000020 OB 1800] ASU ha HARE de sho" is sent to the sentence editing section 107. However, two kinds of files, a voice attribute file with an embedded command and an ordinary file without an embedded command, can also be created. If an ordinary file is additionally created in this way, a voice command will not be a hindrance, and a sentence created by another sentence editing program can be utilized. In a preferred embodiment of the present invention, in response to the settling key being pushed, the unsettled character string is sent not only to the sentence editing section 107 but also to the voice synthesizing section 130. Then, the voice synthesis is performed and the voice adjustment is finally confirmed. Also, the buffer, managed by the kana character string input section 101, is cleared.

Next, when an input key is the other key (step 457), the other process corresponding to the key is per-

formed. For example, when a key for moving a cursor right is pushed, the cursor is moved. When the cursor is moved from the present conversion object character string of an unsettled character string to the character string part of the unsettled character string which is not the present conversion object character string, the conversion object character string is changed to the character string including a character at which the present cursor is located.

Figure 9 is a flowchart showing the control procedure of the voice synthesis control section 131 which received a sentence including an embedded command. If the voice synthesis control section 131 receives a sentence including an embedded command, the section 131 will judge whether the sentence voice command has been embedded in the head of the sentence or not (step 603). In the case where the sentence voice command has been embedded, the voice synthesis control section 131, in accordance with the contents of the voice attribute included in the sentence voice command, instructs the parameter generating section 143 and the voice synthesizer 145 to change parameters and voice synthesis (step 605). In the case where the sentence voice command has not been embedded, the voice synthesis control section 131 next judges whether a character string voice command has been included or not (step 607). In the case where the character string voice command has been embedded, the voice synthesis control section 131, in accordance with the contents of the voice attribute included in the character string voice command, instructs the parameter generating section 143 to generate parameters which correspond to the reading and accent of the character string (step 609). In accordance with the voice attribute included in the command, the voice synthesis control section 131 may also instruct the reading application section 139 and the accent application section 141 to apply "reading" and "accent."

In the case where the character voice command has not been embedded, the input character string is sent to the language analyzing section 133, and the voice synthesis is performed according to a known voice synthesizing method (step 611). The control section 131, in accordance with the contents of the voice attribute included in the sentence voice command, instructs the parameter generating section 143 to generate parameters which correspond to the reading and accent of the character string (step 609).

Thereafter, the next character string is read (step 615) and it is judged whether the character string is the end of a sentence or not (step 617). In the case where the next character string is the end of a sentence, the voice synthesizing process is ended (step 619). In the case where the next character string is not the end of a sentence, the processing is continued and it is judged whether a new character string is a voice command (a sentence voice command or a character string voice command) (step 619). In the case where the new char-

acter string is not a voice command, the character string is sent to the language analyzing section 133.

While the present invention has been described with reference to the embodiment making use of kana-kanji conversion of Japanese, the invention is executable even in the case of other languages such as English. The embedding of the sentence voice command, shown in Figures 3 and 4, is substantially executable with the same contents independently of language. Since such change is a matter which can be easily understood to those having skill in this field, a description is omitted.

The embedding of a character string voice command in language such as English will hereinafter be described. In the case where the present invention is executed for English, the kana-kanji conversion section 103 and the kana-kanji dictionary 105 are not needed. However, when the attribute of an input character string is changed as in the kana-kanji conversion of Japanese, it is also possible to adopt a similar constitution. For example, an input character string is caused to be in an unsettled state, and it is also possible to convert this unsettled character string by an input which instructs font change, a large letter, or a small letter, or by an input which instructs that only the first one character is a large letter. In addition, it is considered that a voice command is embedded into the unsettled character string.

In the case where the present invention is executed in English, a character string, input from a keyboard, is held by the (kana) character input section 101 shown in Figure 2. However, the range of a character string, which has already been input and settled, can be specified by the pointer of a mouse, and the specified range of the character string can be held by the (kana) character string input section 101. The (kana-kanji) conversion control section 113 embeds the voice attribute information held by the voice attribute input section 115 into the information, held by the (kana) character input section 101, in the form of a voice command. The embedding of the voice command is performed in a method similar to the method using the kana-kanji conversion of Japanese.

Figure 10 is a diagram showing an example of a temporary word registration input panel which is displayed to users for adjusting the voice attribute information of a character string voice command. For language such as English, a single word is partitioned off by a delimiter character, and the (kana-kanji) conversion control section 113 can recognize a single word as a single conversion object character string. As with the temporary word registration panel 305 shown in Figure 5, a temporary word registration panel 505 is provided with entries 543 and 547 for adjusting "accent", an entry 545 for adjusting "reading (pronunciation)", and an entry 549 for specifying "a part of speech." User can apply a desired accent and reading to the word "fine" 501 of "It will be fine tomorrow" 503 shown in Figure 10. Therefore, for example, a character string "lead" can be pronounced as "[li:d]" or "[led]." Also, the pronunciation

((led] or [eli:di:]) of "LED" (a light emitting diode) can be changed for each sentence.

According to the present invention, as described above, an embedded command is automatically embedded into an unsettled character string at the time of kana-kanji conversion. Accordingly, the operation is simplified, and furthermore, there is no need for a user to memorize a command itself and there is no mistaken input.

By creating a sentence which includes an embedded command by using both an embedded command valid only to the character and an embedded command valid to the sentence thereafter, it is made possible to change a specific character string only in the sentence and a general dictionary is not influenced. In addition, a fine reading method can be simply defined.

By displaying an embedded-command editing window for a character string unit, it is possible to provide a user interface which is substantially common to ordinary word registration, and the interface is intuitively and easily understandable for users.

At the time of kana-kanji conversion, the voice synthesis of an unsettled character string can be tentatively performed. Therefore, users can confirm the result of the voice synthesis in units of a short character string such as a word. In addition, the operational efficiency is higher than the case where, after a sentence is created, the entire sentence or a character string specified in the document is input to perform voice synthesis, and a voice-command embedded sentence can be created in a short time.

In addition, since there is provided a voice synthesis application which can perform the voice synthesis of a voice command embedded sentence including both an embedded command for a character string and an embedded command for a sentence, voice synthesis adjusted finely by a user can be performed efficiently and effectively.

Claims

1. A method of creating a sentence embedded with a voice command which includes voice attribute information and which is referred to when voice synthesis is performed, the method comprising the steps of:

(a) specifying a character string into which said voice command is embedded;

(b) detecting a user's input which instructs embedding of said voice command into said specified character string;

(c) displaying entries for the user to input voice attribute information of said specified character string; and

(d) embedding a voice command, which includes voice attribute information corresponding to the user's input to said entries, into said specified character string.

2. In a document creating system comprising an input unit, a display, an unconverted character string input section, a character conversion section, a character conversion dictionary, a character conversion control section, a document editing section, and a document storage section, a method of creating a sentence embedded with a voice command which includes voice attribute information and which is referred to when voice synthesis is performed, the method comprising the steps of:

(a) holding an unconverted character string input from said input unit in said character string input section as an unsettled character string;

(b) detecting a user's input, which instructs conversion to a converted character string with respect to said unsettled character string input, from said input unit;

(c) specifying a candidate character string, which is a candidate for a converted character string corresponding to a conversion object character string forming part of said unsettled character string, from said character conversion dictionary in response to the detection of said input which instructs conversion to a converted character string;

(d) displaying said candidate character string on said display;

(e) detecting a user's input, which selects a selected character string which is one of the candidate character strings, from said input unit;

(f) replacing said conversion object character string with said selected character string and taking said selected character string to be a new unsettled character string;

(g) detecting a user's input which instructs embedding of said voice command into said conversion object character string;

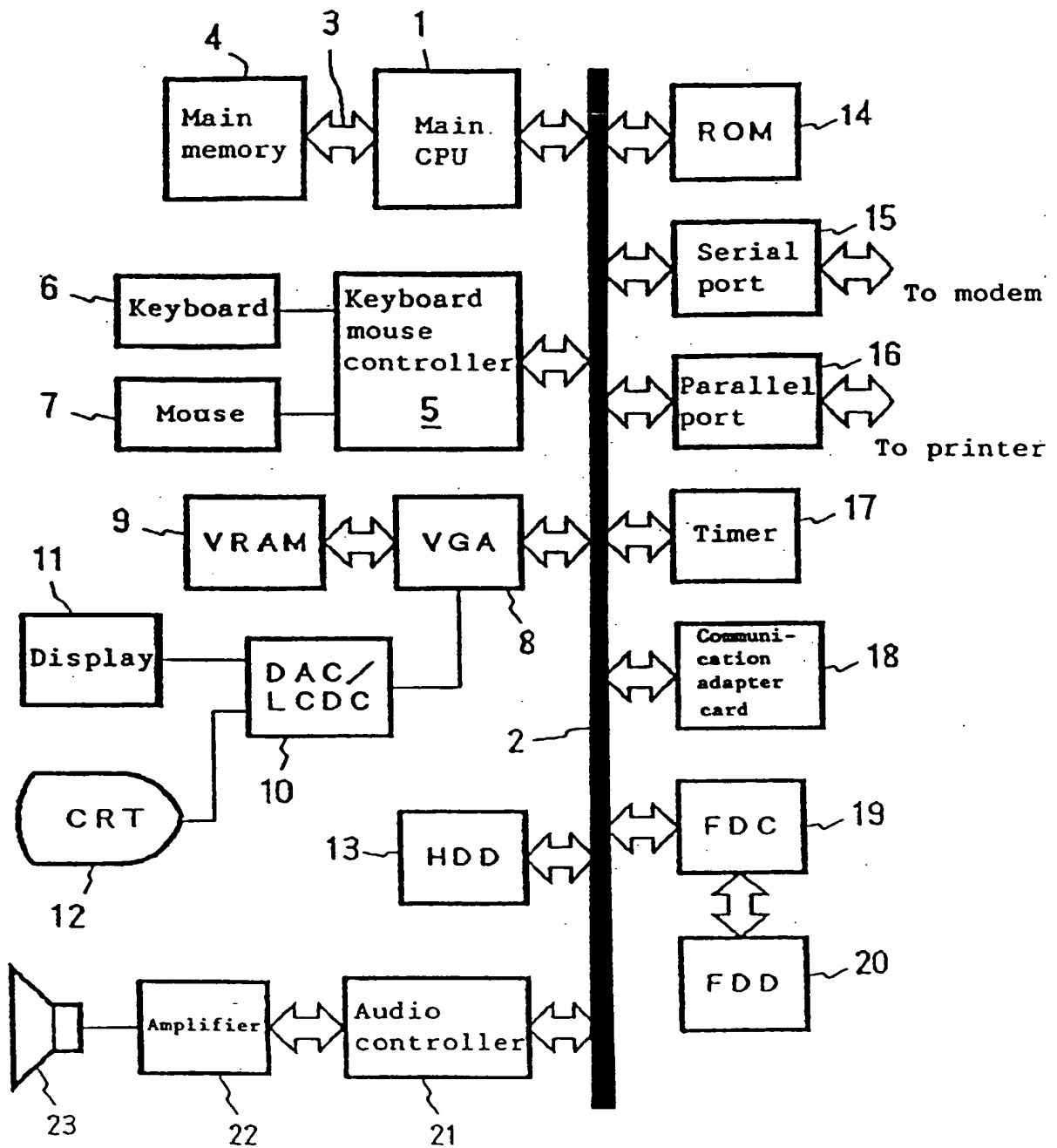
(h) displaying entries for the user to input voice attribute information which includes reading and accent of said conversion object character string which are embedded into said conversion object character string; and

(i) embedding a voice command, which includes voice attribute information correspond-

- ing to the user's input to said entries, into said conversion object character string.
3. The method of Claim 2 further comprising the steps of:
 - (j) detecting a user' input which instructs voice synthesis of said conversion object character string; and
 - (k) performing voice synthesis in accordance with a voice attribute of said voice command.
 4. An apparatus for creating a sentence embedded with a voice command which includes voice attribute information and which is referred to when voice synthesis is performed, the apparatus comprising:
 - (a) an unconverted character string input section for holding a character string input by a user;
 - (b) a character conversion dictionary for managing a converted character string which corresponds to an unconverted character string;
 - (c) a character conversion section for retrieving a candidate for a converted character string which corresponds to said character string held by said unconverted character string input section;
 - (d) a voice attribute input section for holding a voice attribute value adjusted by a user's input; and
 - (e) a character conversion section for instructing said character conversion section to select a converted character string corresponding to the character string held by said unconverted character string input section from said converted character string candidate in response to a user's input and also for embedding said voice attribute value held by said voice attribute input section into the converted character string selected in the form of a voice command.
 5. The apparatus of Claim 4 further comprising:
 - (f) a voice synthesizing section for performing voice synthesis in accordance with voice attribute information embedded in the converted character string embedded with said voice command.
 6. A recording medium for storing a control program which instructs a document creating apparatus to create a sentence embedded with a voice command, said voice command including voice attribute

information and being referred to when voice synthesis is performed, and said control program comprising:

- (a) a program code for instructing said document creating apparatus to specify a character string into which said voice command is embedded;
- (b) a program code for instructing said document creating apparatus to detect a user's input which instructs embedding of said voice command into said specified character string;
- (c) a program code for instructing said document creating apparatus to display entries for the user to input voice attribute information of said specified character string; and
- (d) a program code for instructing said document creating apparatus to embed a voice command, which includes voice attribute information corresponding to the user's input to said entries, into said specified character string.

FIG. 1

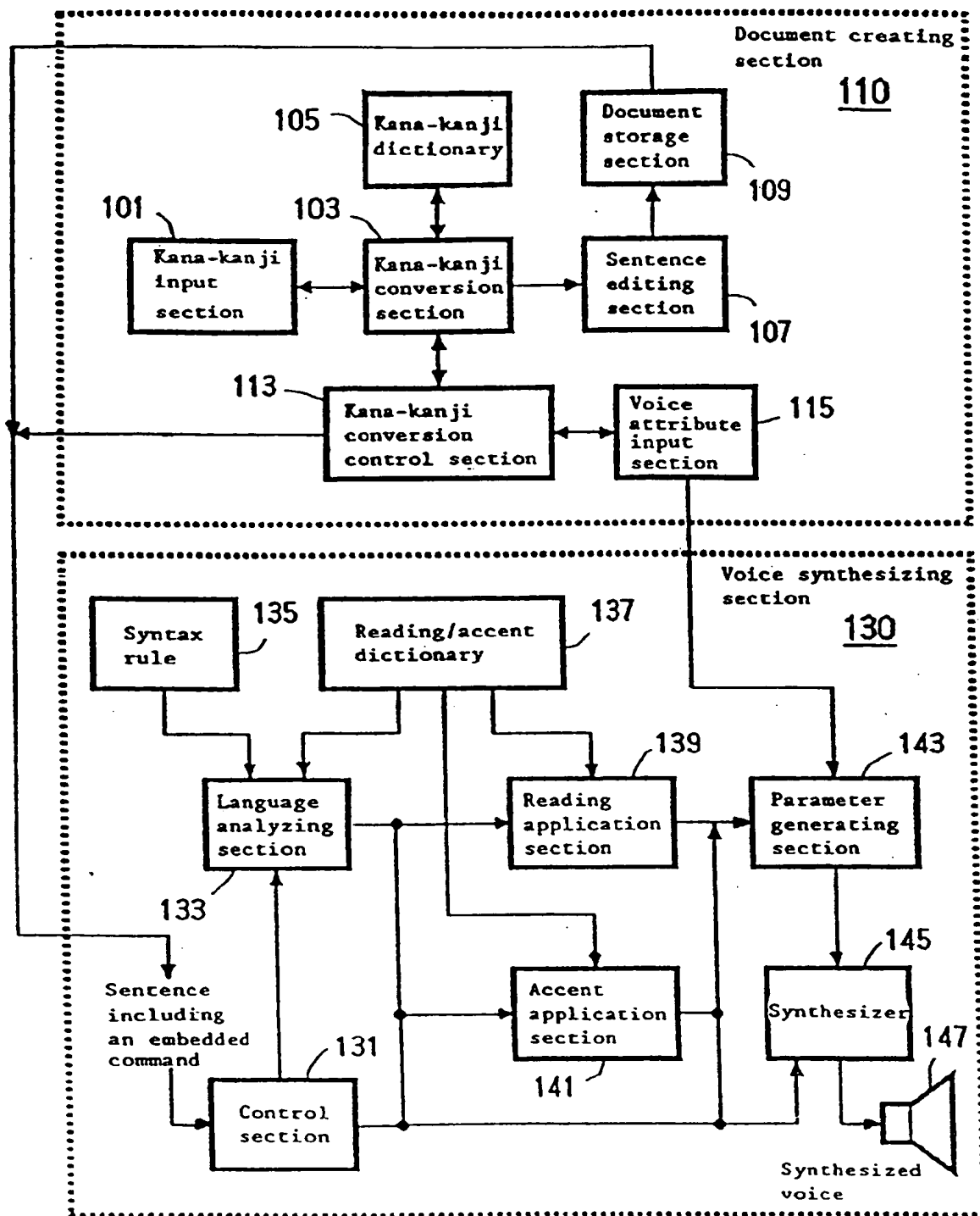


FIG. 2

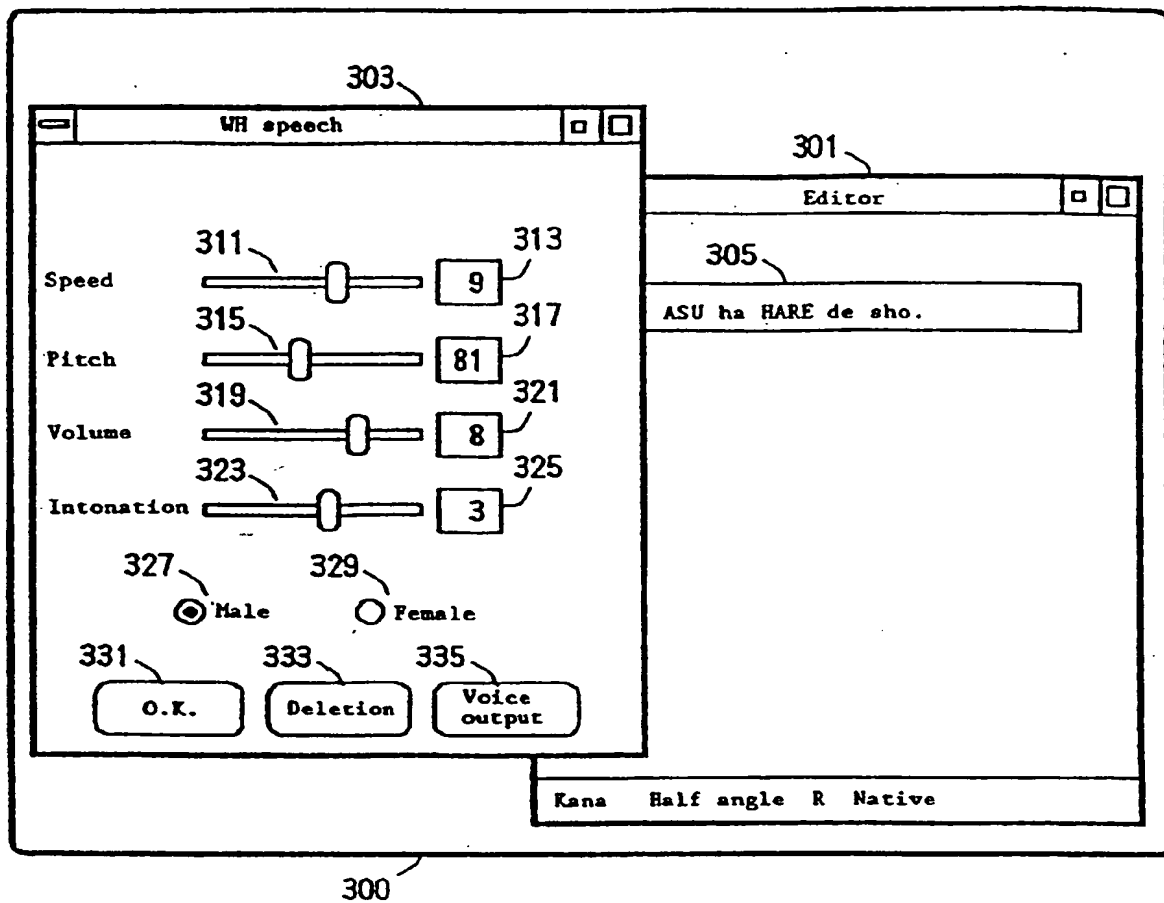


FIG. 3

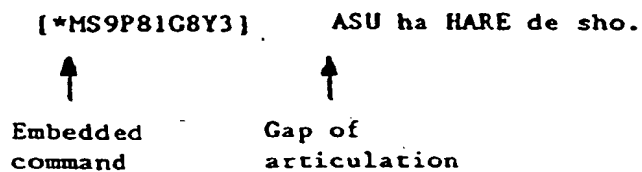
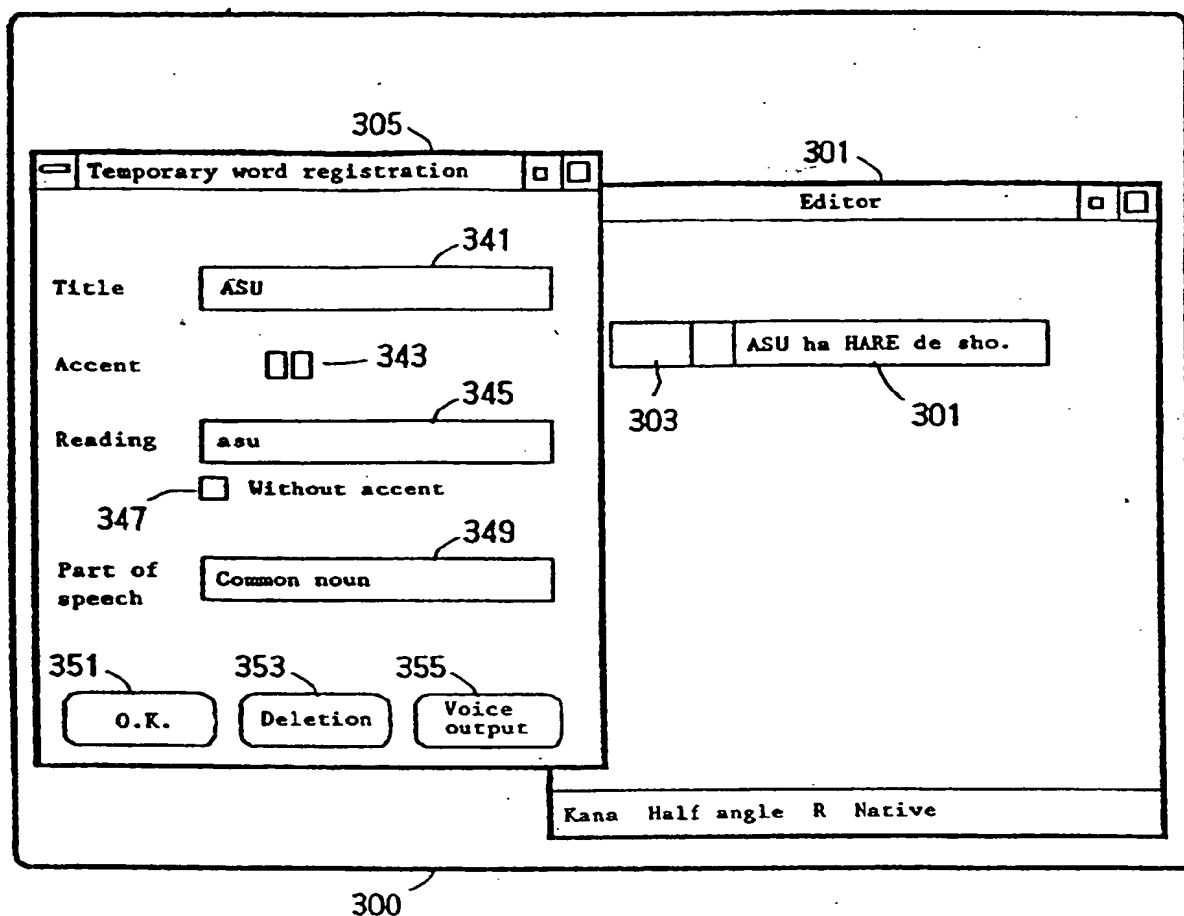
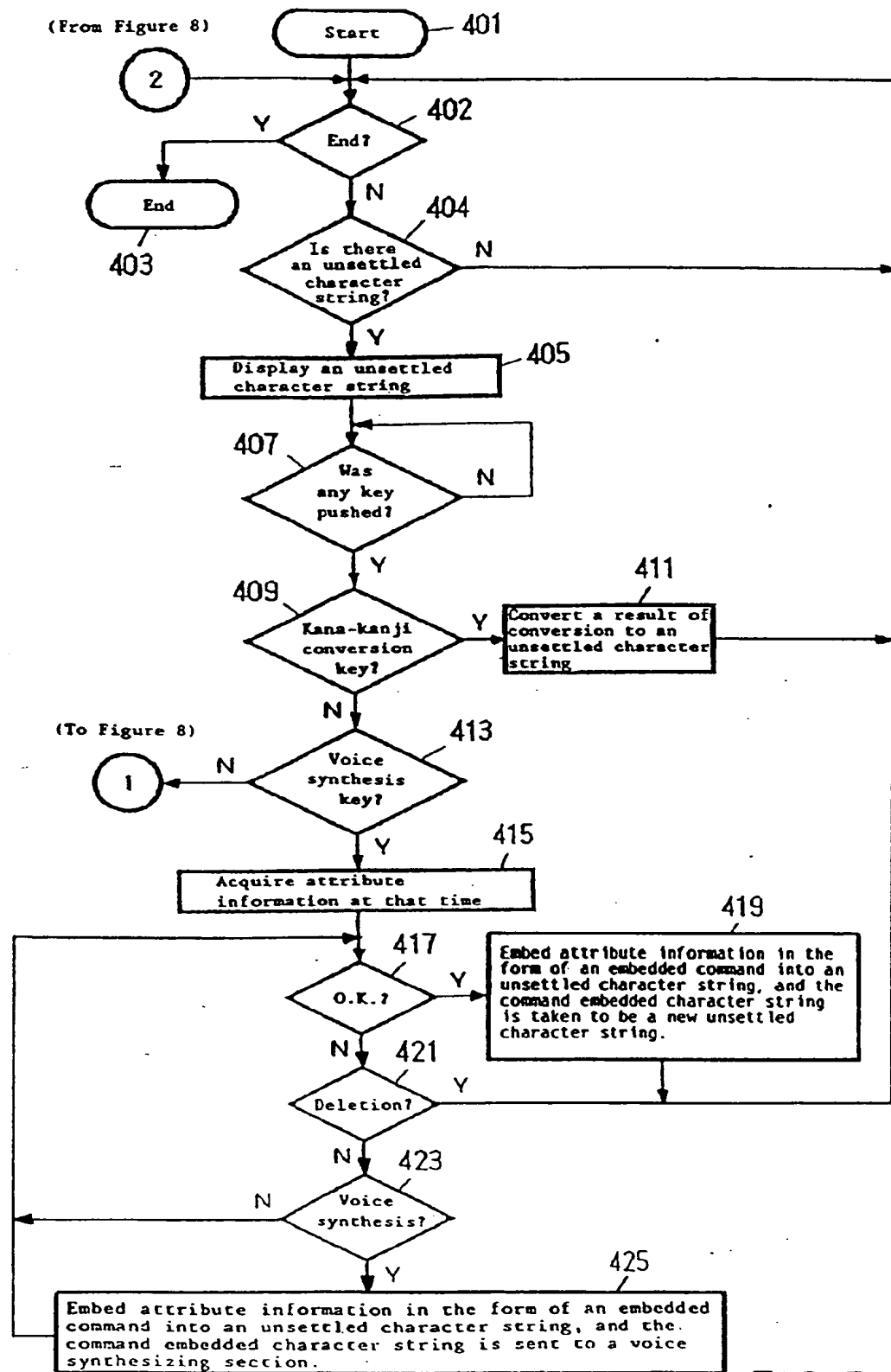


FIG. 4

[*T asu ASU 0 000020 0B 1800] ASU ha HARE de sho.

FIG. 6

FIG. 5



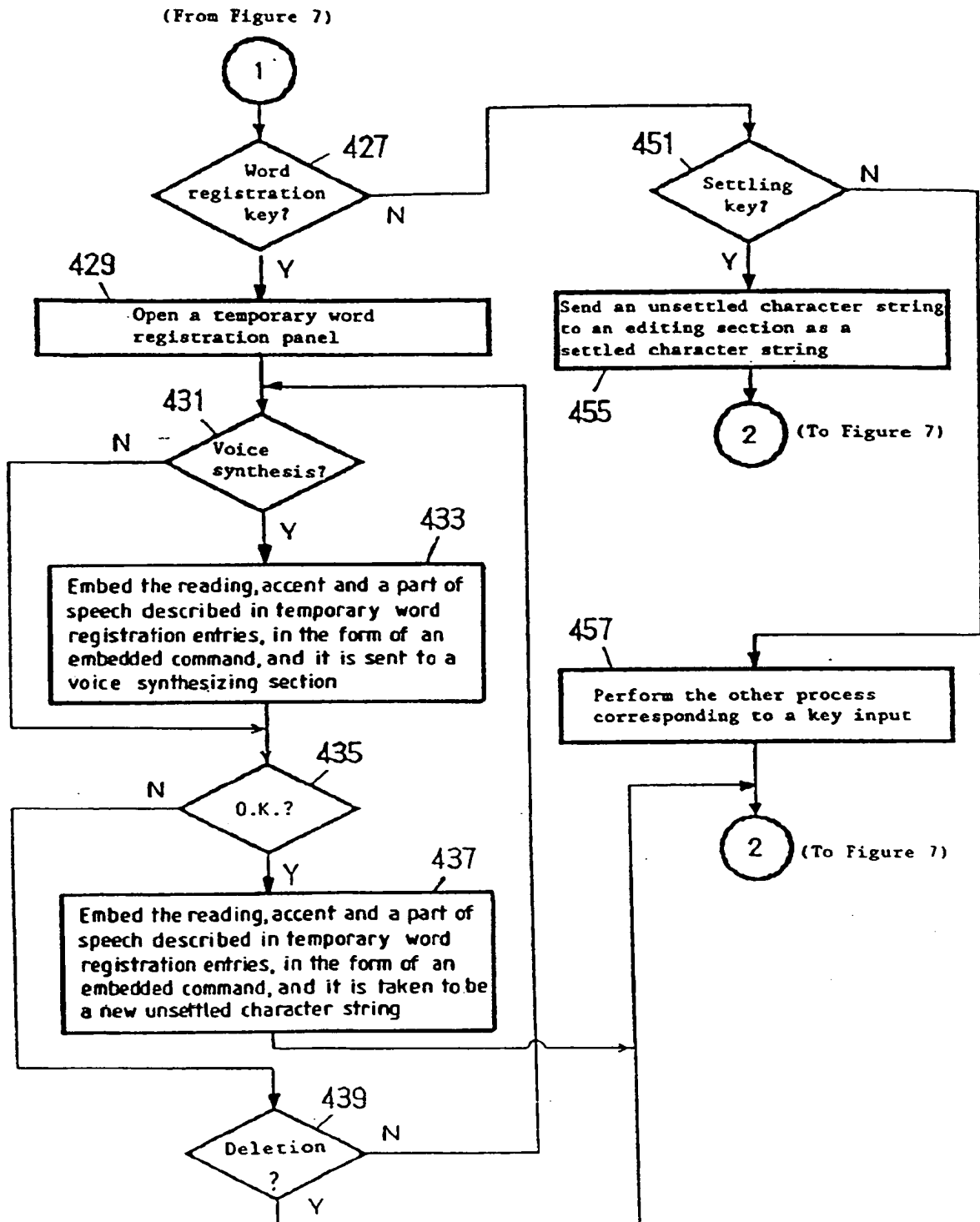


FIG. 8

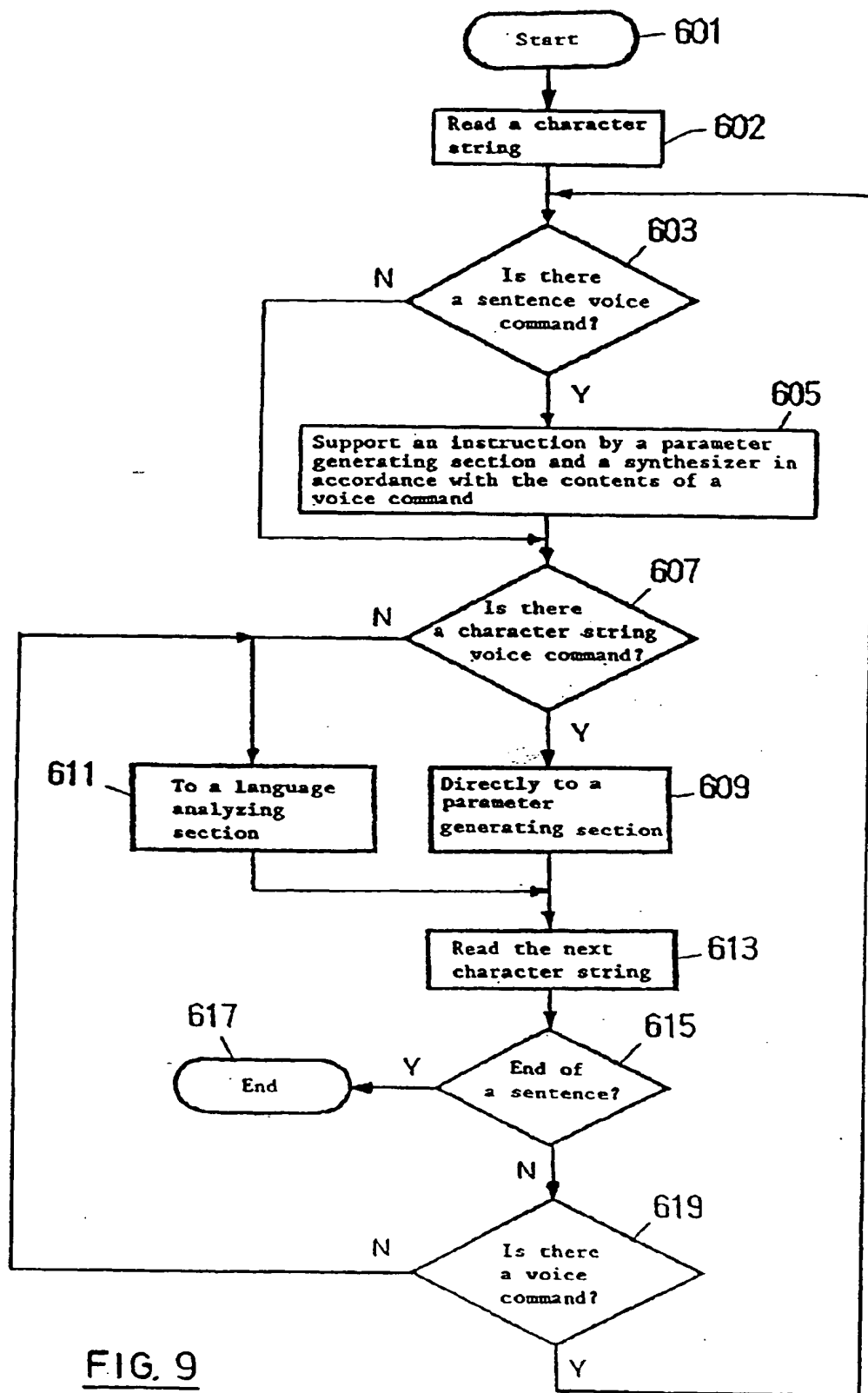


FIG. 9

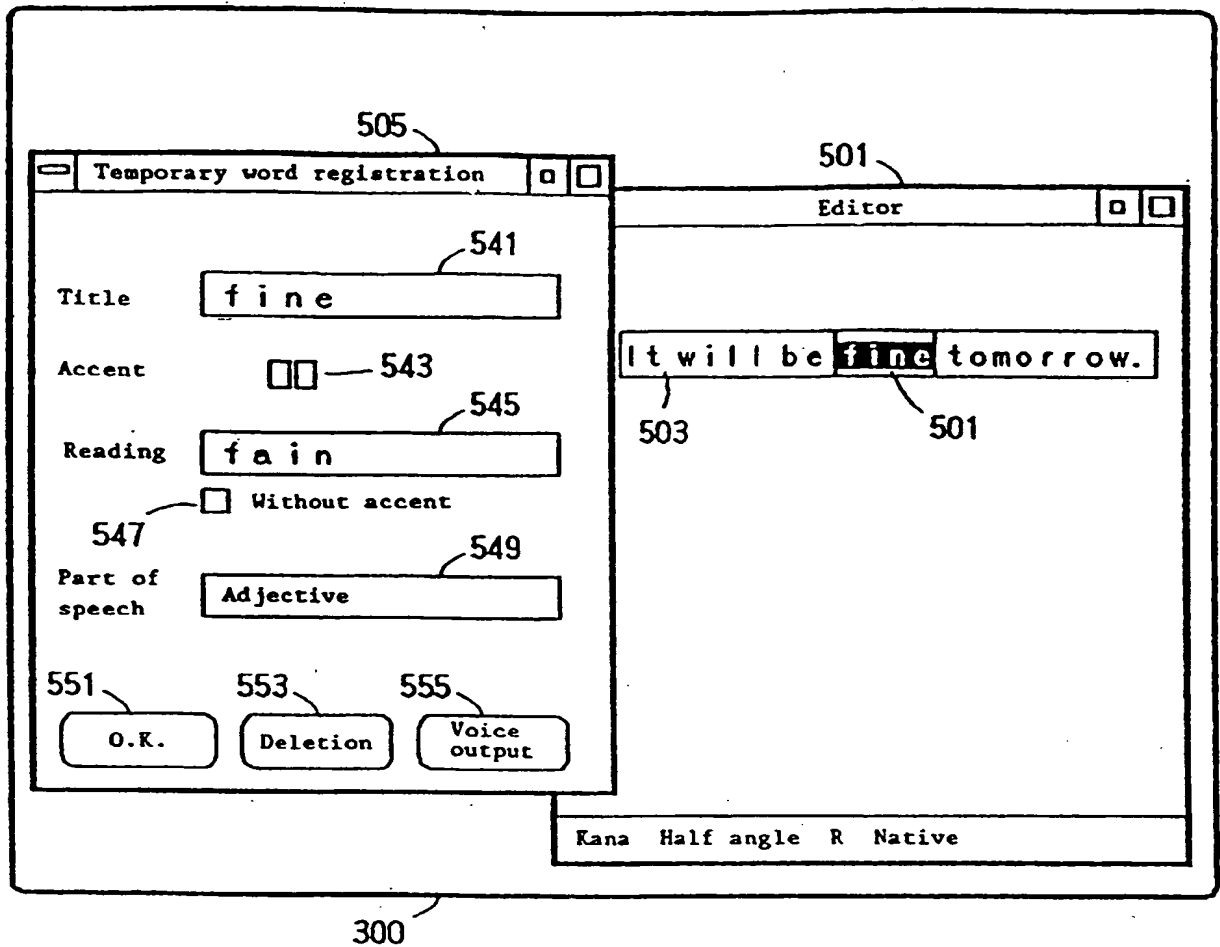


FIG. 10

This Page Blank (uspto)

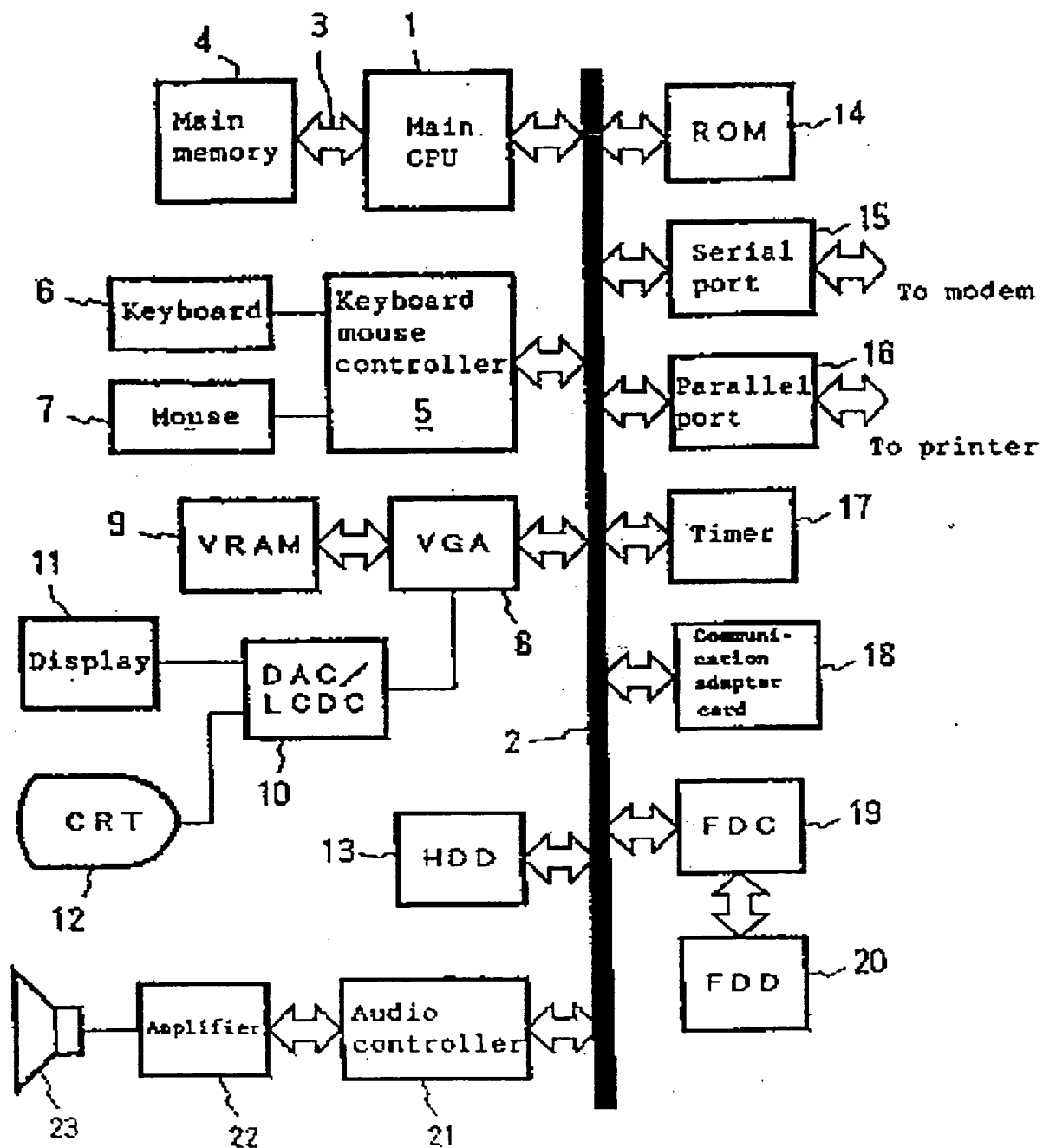


FIG. 1

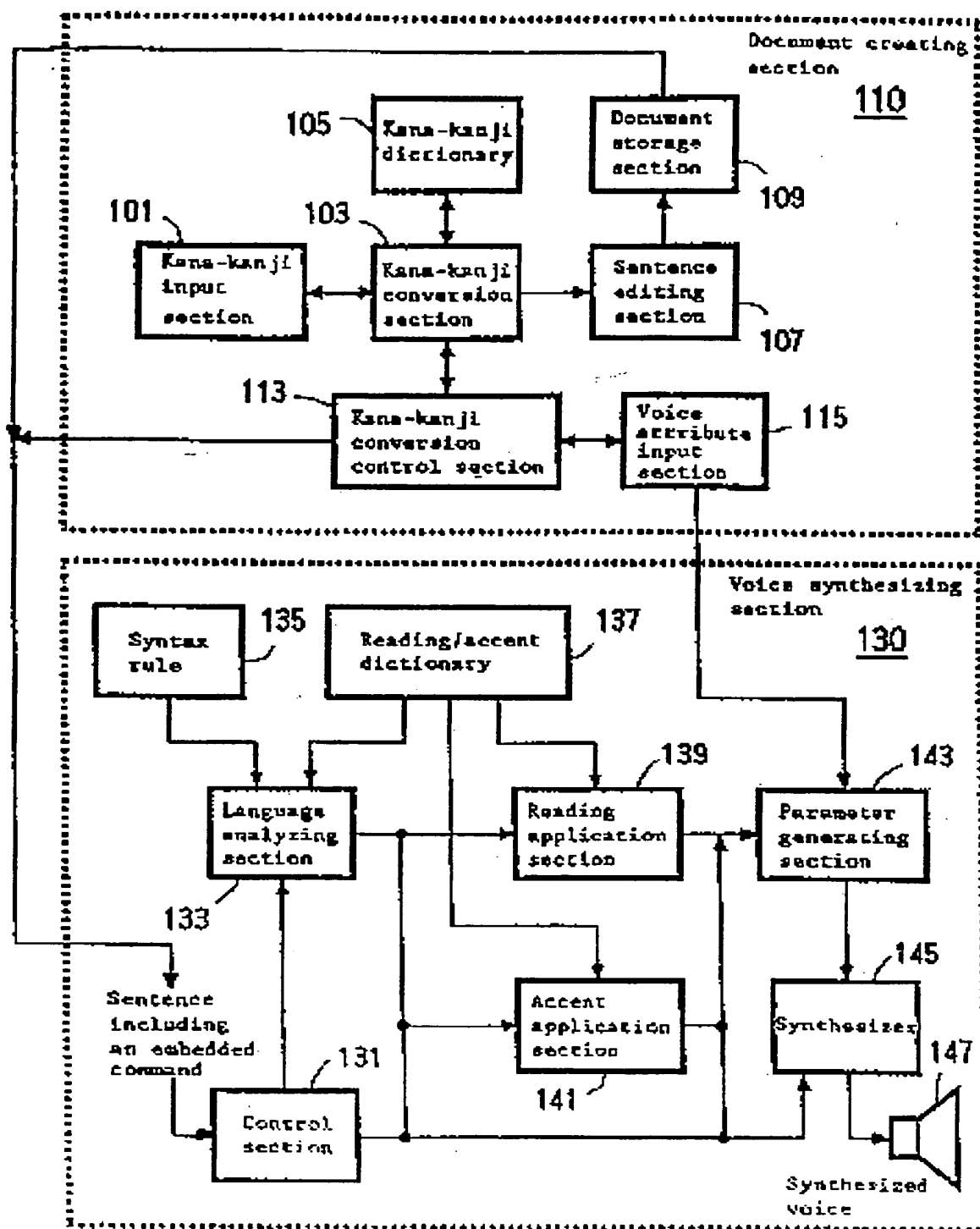


FIG. 2

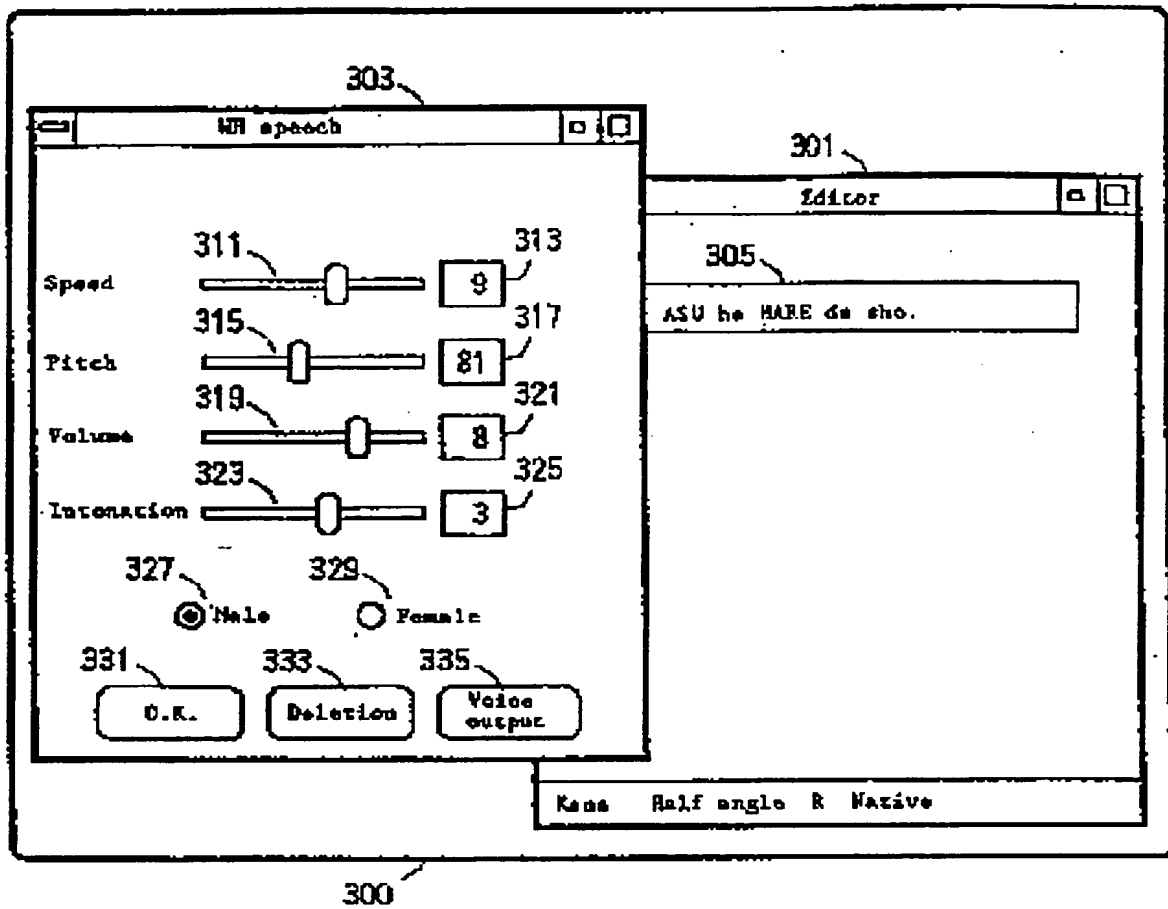


FIG. 3

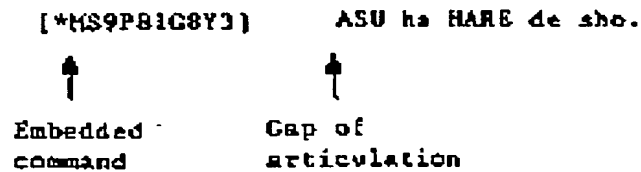


FIG. 4

[*T asu ASU 0 000020 0B 1800] ASU ha HARE de sho.

FIG. 6

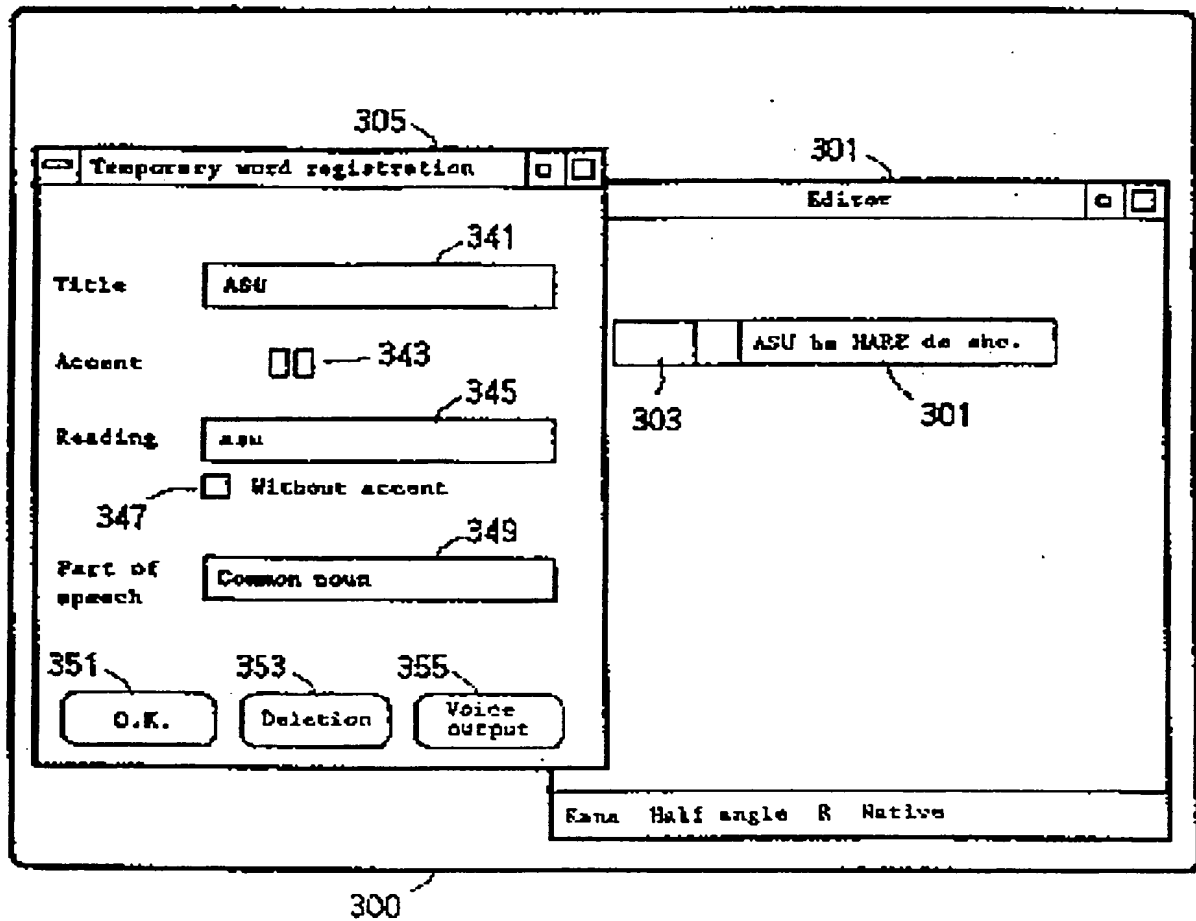
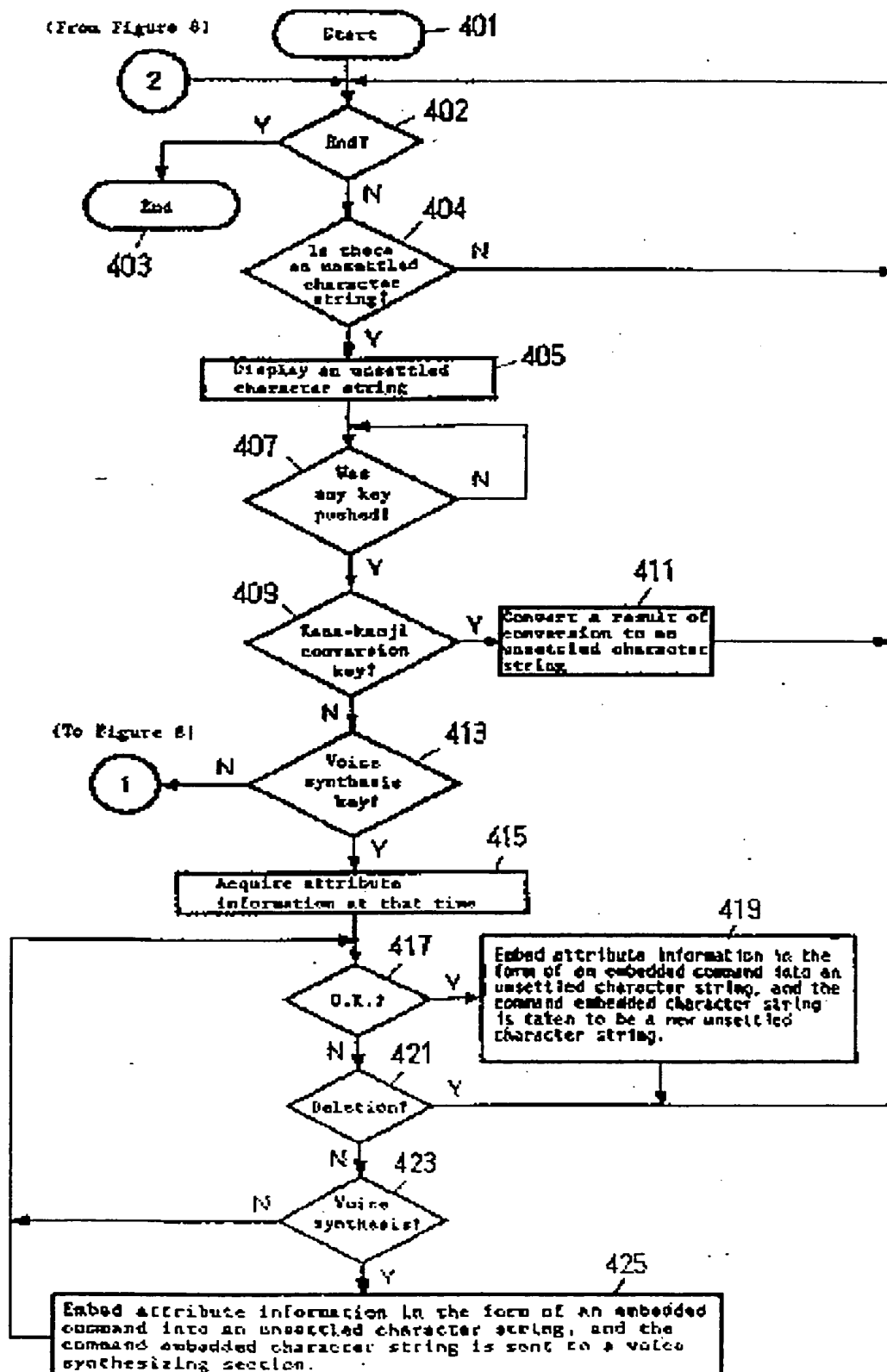


FIG. 5



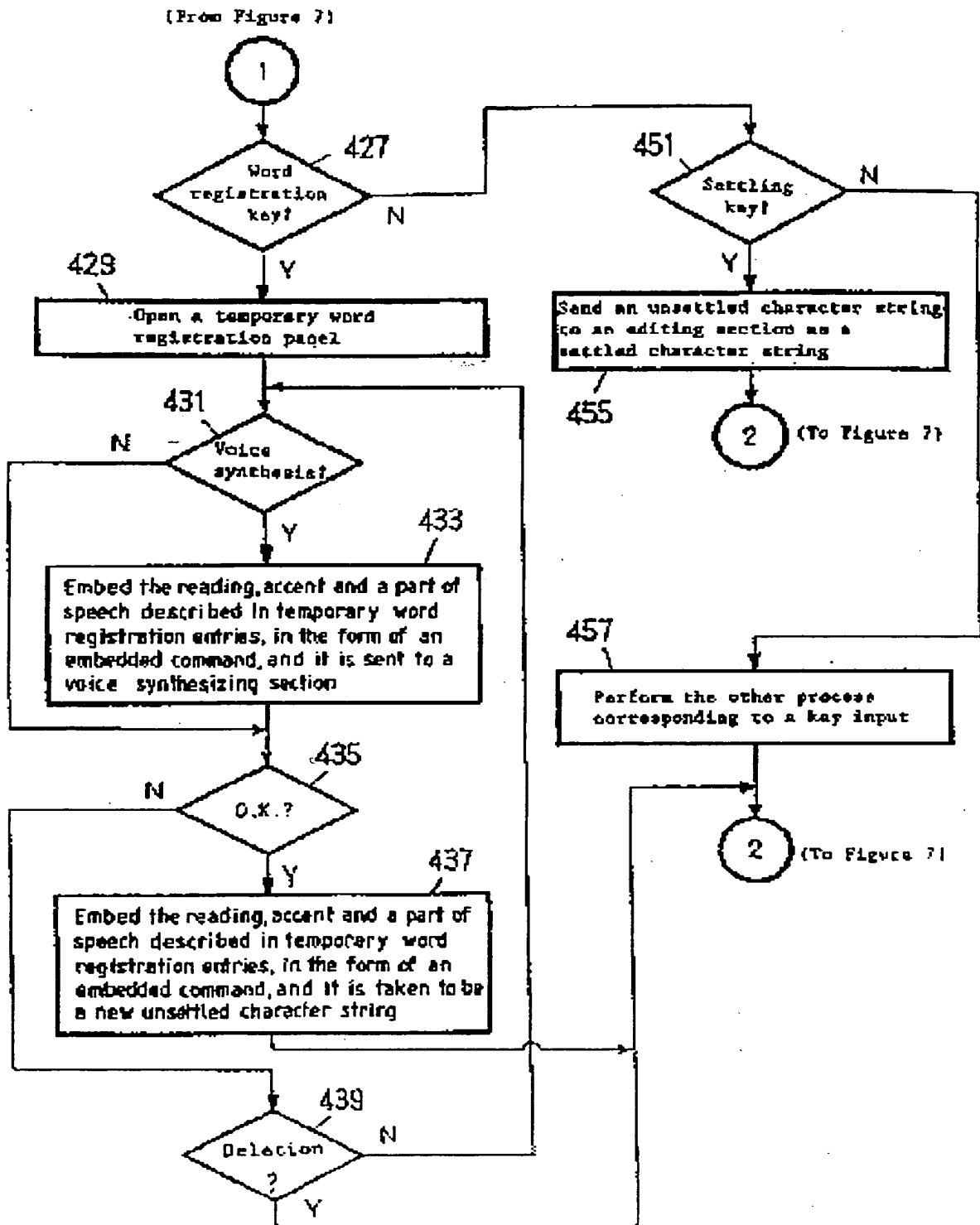


FIG. 8

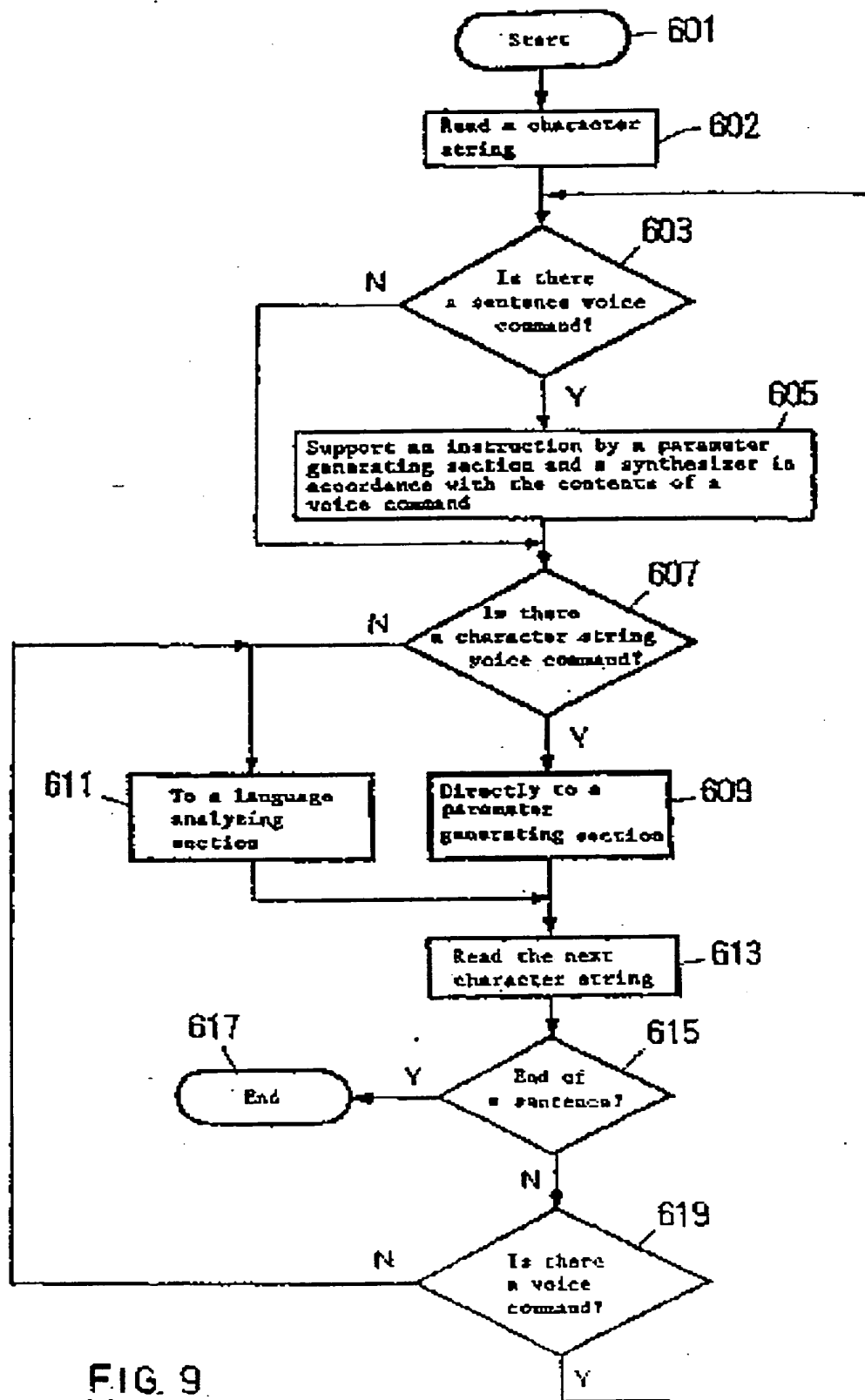


FIG. 9

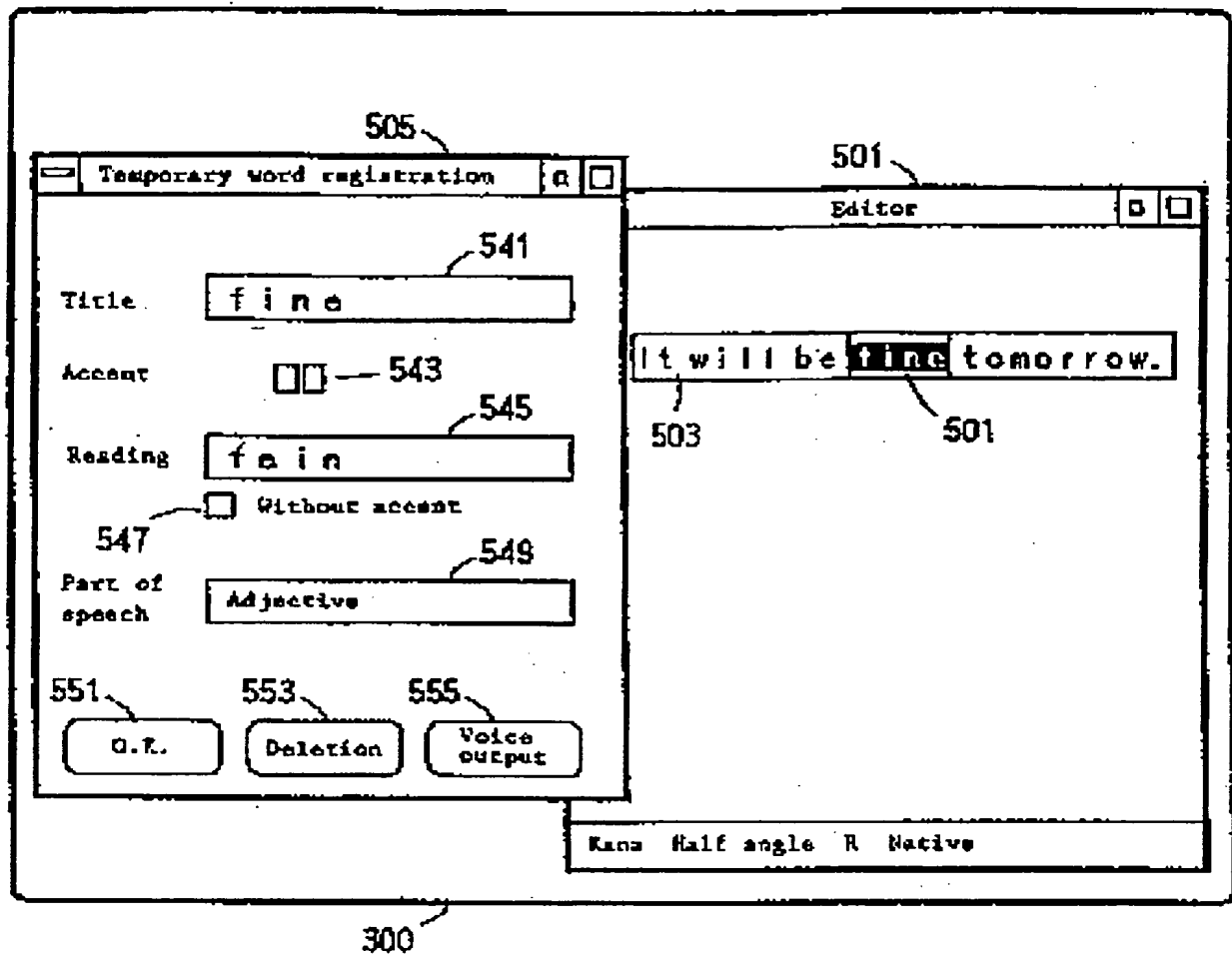
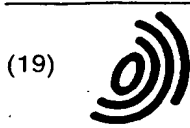


FIG. 10



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 810 582 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
30.09.1998 Bulletin 1998/40

(51) Int Cl.⁶: G10L 5/02

(43) Date of publication A2:
03.12.1997 Bulletin 1997/49

(21) Application number: 97303442.4

(22) Date of filing: 20.05.1997

(84) Designated Contracting States:
DE FR GB

(30) Priority: 30.05.1996 JP 136327/96

(71) Applicant: International Business Machines Corporation
Armonk, N.Y. 10504 (US)

(72) Inventor: Miyashita, Shichiro
Yamato-shi, Kanagawa-ken (JP)

(74) Representative: Zerbi, Guido Maria
Intellectual Property Department,
IBM United Kingdom Ltd.,
Hursley Park
Winchester, Hampshire SO21 2JN (GB)

(54) **Voice synthesizing method, voice synthesizer and apparatus for and method of embodying a voice command into a sentence**

(57) A method which is capable of performing an operation of creating a sentence and an operation of adjusting a voice attribute at the same time.

If a key for embedding an embedding command into an unsettled character string is pushed with the state where the unsettled character string has been displayed after kana-kanji conversion, the voice attribute information held by a voice attribute information input section 115 is embedded in the form of an embedded command into the unsettled character string. Also, if a key for instructing voice synthesis is pushed with this state, voice synthesis is performed according to the embedded voice attribute information.

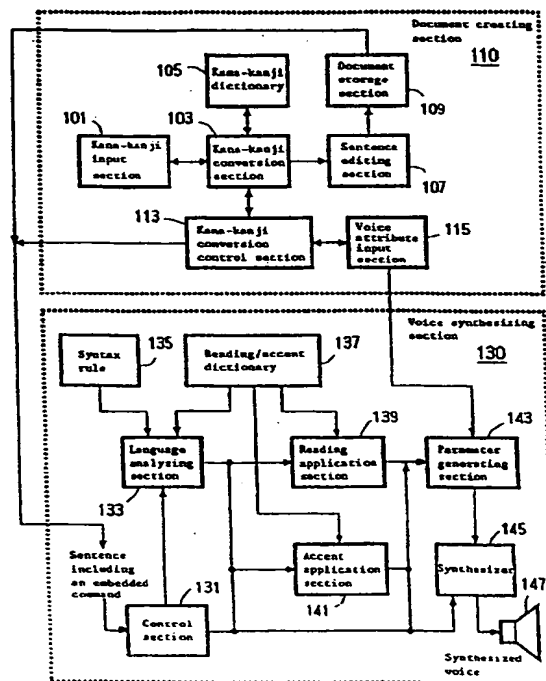


FIG. 2

EP 0 810 582 A3



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 30 3442

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	PATENT ABSTRACTS OF JAPAN vol. 015, no. 067 (P-1167), 18 February 1991 & JP 02 289071 A (RICOH CO LTD), 29 November 1990 * abstract *	1,2,4	G10L5/02
A,D	PATENT ABSTRACTS OF JAPAN vol. 018, no. 516 (P-1806), 28 September 1994 & JP 06 176023 A (TOSHIBA CORP), 24 June 1994 * abstract *	1,2,4	
A	PATENT ABSTRACTS OF JAPAN vol. 017, no. 194 (P-1522), 15 April 1993 & JP 04 342297 A (TOSHIBA CORP), 27 November 1992 * abstract *	1,2,4	
A	KEIICHI NAGAKURA ET AL: "JAPANESE TEXT-TO-SPEECH SYNTHESIZER" REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORIES, vol. 36, no. 5, 1 January 1988, pages 451-457, XP000112299 * paragraph 4.1; figures 1,4 *	1,2,4	TECHNICAL FIELDS SEARCHED (Int.Cl.6) G10L
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 12 August 1998	Examiner Wanzeele, R
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03/92 (Pct/C01)